FLUORESCENT NUCLEOSIDES FOR PROBING PROTON-COUPLED DNA FOLDING IN REAL TIME

Dissertation

zur Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

Vorgelegt der

Mathematish-naturwissenschaftlichen Fakultät

der **Universität Zürich**

von

Guillaume Mata

aus Frankreich

Promotionskomitee

Prof. Dr. Nathan W. Luedtke (Vorsitz)

Prof. Dr. Cristina Nevado

Prof. Dr. John A. Robinson

Prof. Dr. Roland K. O. Sigel

Zürich, 2015

Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich im Frühjahrssemester 2015 als Dissertation angenommen.

Promotionskomitee: Prof. Dr. Nathan W. Luedtke (Vorsitz)
Prof. Dr. Cristina Nevado
Prof. Dr. John A. Robinson
Prof. Dr. Roland K. O. Sigel

Dedicated to my parents,

to Anne-Catherine,

Acknowledgments

The accomplishment of this work was realized with the participation of many people who supported me along this whole adventure. The following ladies and gentleman are profoundly acknowledged.

At first, I would like to express my deepest gratitude to Prof. Dr. Nathan W. Luedtke for his innovative ideas and outstanding perception of research in organic and bioorganic chemistry. Thank you for your constant support and availability throughout this entire thesis and for giving me all the independence I needed. I am very grateful to Prof. Dr. Cristina Nevado, Prof. Dr. John A. Robinson and Prof. Dr. Roland K. O. Sigel for serving on my thesis committee and for our good working relationship from the beginning to the end of this thesis. I would like to thank Prof. Dr. Robert Häner who kindly accepted to give his expert review on the thesis. I am also thankful to Prof. Dr. Gilles Gasser for our friendly scientific discussions. I am very grateful to Dr. Ian M. Eggleston, who was a great supervisor during my Master thesis at the University of Bath, who sent great recommendation letters since many years and contributed to all the positions and grants I obtained.

I would like to thank our research group, Dr. Jawad Alzeer, Marco Brandstätter, Dr. Anaëlle Dumas, Dr. Sarah Hentschel, Sabrina Huber, Theodor Marsoner, Paul McEoin, Alessandra Messikommer, Dr. Anu Naik, Dr. Anne Neef, David Oesch, Johanna Penner, Dr. Ulrike Rieder, Fabian Roth, Olivia Schmidt, Dr. Martin Seyfried, Therese Triemer, Dr. Bala Vummidi and Barbara Zampoli for the great atmosphere inside the laboratory and outside the University. Many thanks as well to Dr. Denise Pauli, Basil Lörtscher and Victor Samper for all the fun we had together during lunch breaks. I would like to emphasize my acknowledgements to David Oesch, who was a great labmate since we have met. Thank you very much for all discussions about chemistry, snowboarding and Swiss mountains. I wish to give special thanks to Anaëlle Dumas who started very successfully our common research topic. Thank you for sharing your precious advises during this thesis, your contagious motivation for research and your great friendship. I also would like to further thank Ulrike Rieder for sharing her successful experience in DNA and kinetic experiments and Therese Triemer and David Oesch for proofreading the Zusammenfassung.

Acknowledgments

I am very grateful to the Swiss National Science Foundation for generous financial support, the University of Zürich for providing outstanding facilities and the Swiss Chemical Society for travel awards. I wish to thank the Graduate School of Chemical and Molecular Sciences and Roland Sigel for their first invitation to Zürich and financial support during this thesis. Furthermore, I would like to acknowledge the mass spectrometry team at the University of Zürich, including Urs Stalder, Dr. Jean-Christophe Prost, Rahel Bucher and Dr. Laurent Bigler for their participation to many analytical measurements. I am especially thankful to Jean-Christophe Prost for his friendship and authenticity.

I wish to thank my family and especially my parents, Marie-Laure and Horacio Mata as well as my sisters, Caroline and Aurélie, who supported me from the beginning of this adventure and helped me to realize my ambitions.

This thesis was only made possible with the wonderful support of Anne-Catherine Remond. Thank you for your constant encouragement and trust. I feel very lucky to have met such a generous and caring person.

Table of Contents

Chapter 1 | Introduction: DNA is The Central Molecule of Life

Chapter 1

DNA is The Central Molecule of Life

1.1. The Double Helix

The transfer of genetic information from nucleotide sequences to proteins was one of the most important discoveries in science.[1] It all started with the discovery of DNA in 1869, when the Swiss scientist Friedrich Miescher published his theories about how biological molecules could encode genetic information.[2] After 80 years of research by many groups,[3-6] Francis Crick and James Watson proposed the double helix as an important secondary structure of DNA.[7-8] In this same year, the Hershey–Chase "blender experiment" definitively proved DNA as being carrier of the genetic code.[9] Duplex DNA, however, can be considered a "resting state" of genetic code, where information actively flows into and out of single-stranded DNA.

1.1.1. Biological Functions of DNA

For many years, DNA has been regarded by most as a uniform double helix and a passive library of genetic information. However, DNA structures are highly dynamic and their functions are potentially diverse.[10] DNA can exist as a double helix or single-stranded DNAs that fold into alternative conformations, such as hairpin, G-quadruplex[11] and *i*-motif[12] structures. One of the most important functions of DNA is the coding of proteins. The amount of coding sequences represent, however, less than 2% of the human genome.[13] The roles of the remaining 98% of DNA sequence are still under study, but highly conserved sequences found throughout the genome are likely to be functional.[14] Studies focused on deciphering the roles of transient structures, such as G-quadruplex and *i*-motif DNA have provided growing support that these structures are biologically relevant.[15]

1.1.2. Structural Features of Duplex DNA

DNA is a polymer of nucleotides that are composed of a heterocyclic nucleobase and a pentose carbohydrate linked *via* a phosphate backbone. The heterocyclic nucleobases are divided in purines and pyrimidines that are present in equal amounts in double-stranded DNA. Adenine (**A**) pairs with thymine (**T**) through two hydrogen bonds, and guanine (**G**) pairs with cytosine (**C**) through three hydrogen bonds (**Figure 1.1**). The very similar geometries of the **A·T** and **G·C** base pairs contribute to the fidelity of replicative polymerases.[16] The arrangement of base pairs is highly flexible and can be described by helical parameters, such as propeller twist and buckle within a single base pair, and shift and slide between

two stacked base pairs. The preferred conformation of the nucleotides is *anti* where the hydrogen bonding face of each nucleobase is pointing away from the carbohydrate. At physiological pH, the negatively charged phosphate backbone is projected toward the polar environment, whereas the heterocyclic nucleobases are pointing toward each other to create a hydrophobic cavity.

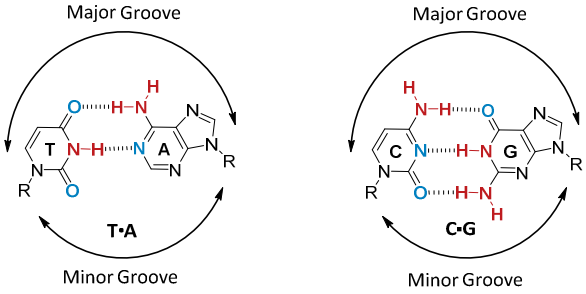


Figure 1.1. Watson-Crick hydrogen base pairs of **T·A** and **C·G** in duplex DNA. Hydrogen bond donors are shown in red and hydrogen bond acceptor are shown in blue.

A- and B-form helicies consists of two anti-parallel strands that are twisted to give a right-handed double helix containing a major and minor groove (**Figure 1.2**). B-DNA has a helical turn of 3.4 nm and a diameter of 2.0 nm, whereas A-DNA has a more compact (2.5 nm) and wider (2.3 nm) structure that is usually observed under dehydrating conditions.[17]

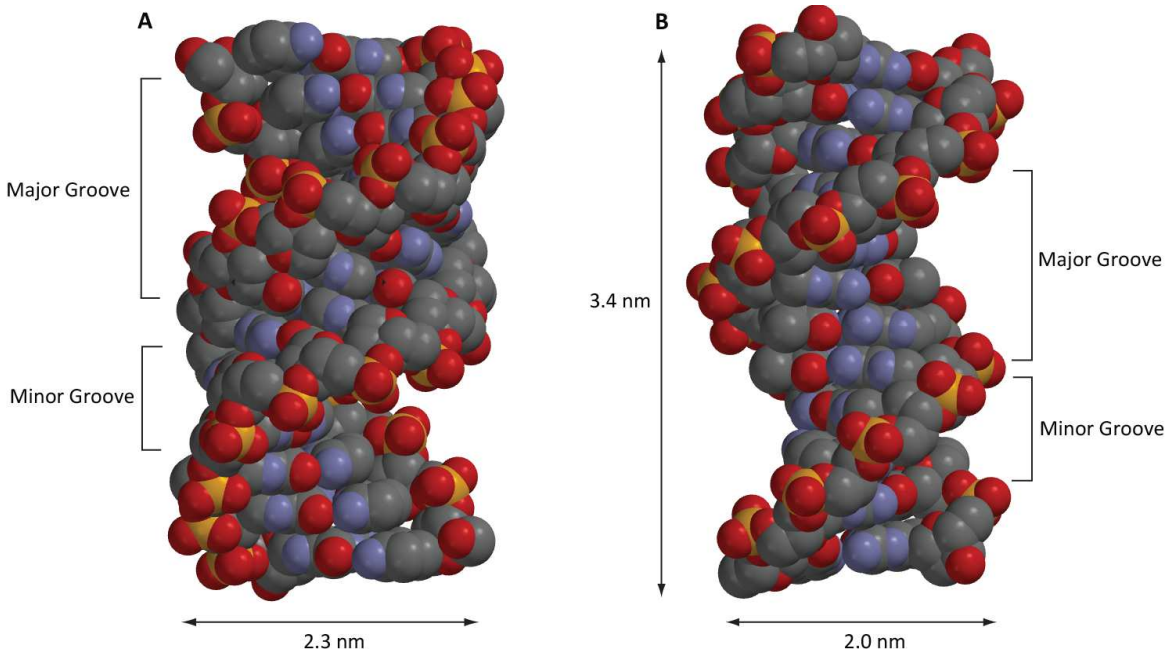


Figure 1.2. Structural representations of (A) A-DNA and (B) B-DNA double helicies. 3D structures are shown as space filling models obtained from PDB files 3V9D[18] and 2BNA,[19] respectively.

1.2. The *i*-Motif DNA

1.2.1. Structural Features of *i*-motif

Four-stranded *i*-motif DNA structures are composed of two parallel-stranded duplexes zipped together in an anti-parallel orientation through intercalated dimers of cytosine, giving $C{\cdot}C^+$ base pairs (**Figure 1.3**). Since this tetrameric DNA features a unique architecture of intercalated base pairs, it was called the *i*-motif.[20] *i*-Motif can form in an intermolecular fashion from multiple stands or in an intramolecular fashion from a single strand. The resulting structures differ in the number of base pairs and loop topologies.[21] Intramolecular *i*-motif structures can adopt different conformations and intercalation topologies with some nucleosides adopting an *anti* glycosidic bond conformation.[22-23] The spatial arrangement of proton-bound dimers of cytosine results in two broad and flat major grooves and two extremely narrow minor grooves (**Figure 1.3A** and **B**). The structure of the proton-bound dimers of cytosine involves the formation of three hydrogen bonds between one *N*3-protonated and one neutral cytosine. The position of the proton is asymmetric, with a double-well potential.[24] Two stacked base pairs belong to two distinct duplexes and are perpendicular to one another (**Figure 1.3C** and **D**). The degree of freedom of the base pairs in *i*-motif DNA is relatively limited. Although, propeller twist and buckling are allowed within each base pair *i*-motifs exclude shifts, rolls and twisting of the stacked base pairs.[20] The helical twist between adjacent $C{\cdot}C^+$ pairs (12-17°) is much smaller than B-DNA (36°) or A-DNA (32.7°). The lifetime of the proton bound dimers are very short for the outermost pairs (lifetime ≈ 1 ms) and relatively long for the innermost pairs (≈ 1 hour),[25] which suggest some dynamics from the residues adjacent to the loop regions (**Figure 1.3E**). X-ray crystallographic analyses demonstrated that *i*-motifs have high densities of nucleobases, due to compact interactions between four-stranded segments, as well as $C{\cdot}H^+$-C base pairs that exhibit an unusually close base-stacking distances (3.1 Å) as compared to duplex DNA (3.4 Å).[26-27]

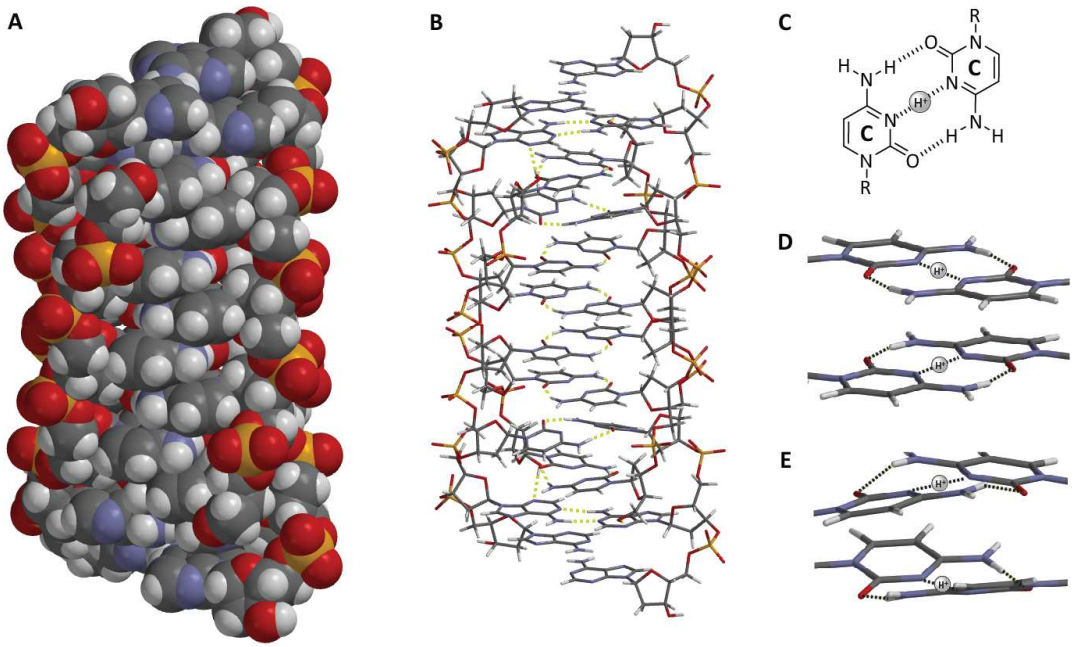


Figure 1.3. Structural representation of *i*-motif DNA. 3D structure of *i*-motif DNA as a (A) space filling and as a (B) tube representation. The structure was obtained from PDB 1YBR[25] and viewed with Spartan 10. (C) 2D schematic representation of the intercalated cytosine base pair. (D) 3D structure of intercalated cytosine base pairs from (D) the innermost part and (E) the outermost part.

An important feature of *i*-motif DNA is the absence of aromatic stacking interactions between adjacent base pairs. The only stacking interactions involve the exocyclic carbonyl and amino groups. This is, however, compensated by the presence of an intramolecular C–H··O hydrogen bonding network between the furanose moieties of the antiparallel backbones.[28] In vacuum, the three hydrogen bonds in the C·C$^+$ base pair are predicted to be energetically more important (169.7 kJ mol^{-1}) as compared to the Watson-Crick C·G (96.6 kJ mol^{-1}) and neutral C·C base pairs (68.0 kJ mol^{-1}).[29] In addition, the association of the four strands facilitate electrostatic interactions of the protonated C·C$^+$ base pairs and the negatively charged phosphate backbone. The stability of *i*-motif DNA is therefore strongly dependent on external factors such as ionic strength, temperature, pH and hydration.[30] Since the protonation of the cytosine residues constitutes an absolute requirements for folding, the pH plays a crucial role in the stability of the structure. Cytidine is the most intrinsically basic canonical nucleobase, with a p*K*a = 4.5 for its conjugate acid in water.[31-32] The maximal stability of the *i*-motif occurs at pH values near the p*K*a of cytosine, but long-range electrostatic and local hydrogen bonding interactions can raise the effective p*K*a (p*K*) of protonated cytidine residues in folded structures.[33] The apparent p*K* values for cytosine residues

in *i*-motif structures are typically in the range of 5.5 – 6.6 *in vitro*,[34] but can reach values above p*K* ≈ 7, especially under molecular crowding conditions.[35-38]

1.2.2. Evidence for *i*-motif *in vivo*

The folding of intramolecular *i*-motif structures raises the apparent p*K*a value of cytosine from p*K*a = 4.5 to p*K* = 5.5 – 6.6 *in vitro*, but the presence of slightly acidic conditions seems necessary for the formation of stable structures.[34] Although the extracellular pH of normal eukaryotic cells is typically 7.4,[39] cancer cells with dysregulated pH have a lower extracellular pH in the range of 6.7 – 7.1.[40] In these conditions, the protonation of cytosine residues would be more favorable and *i*-motif DNA structures would be more likely. The crowding nature of the *in vivo* environment could also increase the marginal stability of *i*-motif DNA under physiological conditions, as shown in the *in vitro* studies where *i*-motif structures were folded at pH as high as 8.0.[35-36]

Cytosine rich sequences $(C_3N_{1-7})_4$ are present throughout the human genome, and putative *i*-motif-forming sequences have been found in more than 40% of gene promoter regions.[41] Stable *i*-motif structures have been reported in several oncogenes, such as *VEGF*,[42] *BCL2*,[43] *RET*,[44] c-*MYC*[45] and *Rb*.[46] The unique geometry and charge distribution of *i*-motif DNAs represent an interesting target for specific binding.[47] Indeed, a number of proteins have been found to specifically bind to cytosine-rich sequences *in vitro* at neutral pH.[48-49] However, the formation of *i*-motif structures was not demonstrated in these studies. Recently, it has been shown that the highly dynamic *i*-motif structure in equilibrium with the corresponding hairpin could be stabilized by small molecule binding.[50-51] This resulted in the significant upregulation of the *BCL2* gene and provided strong evidence that these DNA structures are biologically relevant.[50-51]

1.3. Telomeric DNA

Human telomeres are nucleoprotein complexes which protect chromosomes from degradation. Telomeric DNA contains the repeated sequence $(^{5'}GGGTTA^{3'})_n$ that is paired with its complement strand, except for a short (24 – 400 bases) single-stranded 3'-overhang. The complementary repeated sequence $(^{5'}CCCTAA^{3'})_n$ which is known to fold into an *i*-motif under acidic conditions, has been

reported to exist transiently as a single-stranded 5'-overhang (**Figure 1.4**).[52] Telomeric DNA gets shortened by each successive round of cellular replication, possibly resulting in a molecular clock that limits the total number of replicative cycles a normal cell can undergo.[53] To overcome this "end replication problem," most cancer tissues maintain genomic stability by telomerase-mediated extension of telomeric DNA.[54] The stabilization of G-quadruplexes structures in the telomeres has been shown to inhibit telomerase activity.[55-57] Since telomerase is expressed in the majority of malignant tumor cells and in relatively few somatic cells, it was recognized as a potential cancer-specific target.[58]

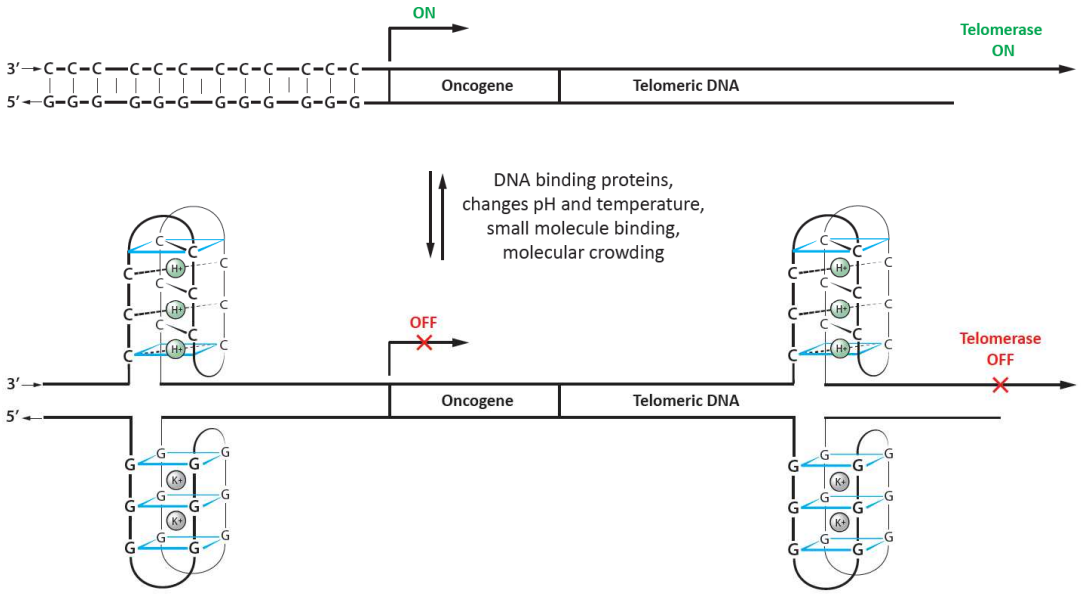


Figure 1.4. Structural representation of the equilibrium between duplex and quadruplex structures in the control of oncogene transcription (upper) and telomerase inhibition (lower).

Due to the self-complementarity nature of duplex DNA, G-quadruplex and *i*-motif structures can exist in a dynamic equilibrium with the Watson-Crick double-stranded DNA.[59] The double helix is more stable than the corresponding single-stranded DNAs folded into intramolecular structures at physiological conditions of pH and temperature. The large kinetic barrier for duplex dissociation can be reduced during transcription and replication processes where negative supercoiling can facilitate local unwinding of the double helix. Intramolecular DNA structures could be favored under these conditions. The combination of multiple factors such as small molecule binding,[60] chaperone,[61] molecular crowding,[37] and negative supercoiling[45,62] could be responsible for shifting the equilibrium toward intramolecular structures (**Figure 1.4**). While mounting evidence supports the presence of G-quadruplex

structures,[15] very little is known about the potential biological relevance of the equilibrium between duplex DNA and G-quadruplex / *i*-motif structures *in vivo*. Taken together, this strongly encourages the development of a sensitive method for the direct detection of *i*-motif and G-quadruplex DNA structures.

1.4. The G-quadruplex DNA

1.4.1. Structural Features of G-quadruplex

G-quadruplex structures are characterized by the presence of two or more stacked "G-tetrads" each containing four guanine residues involved in Hoogsteen hydrogen bonds (**Figure 1.5**).[63-64] G-quadruplexes can be assembled from one, two, three or four separate strands and can display various topologies according to different factors, such as sequence, loop size, salts in solution, temperature and pH.[65] The core structure delimits four negatively charged grooves linked together by three types of loops, called lateral, diagonal or propeller.[66] All parallel G-quadruplexes contain guanosine residues with *anti* glycosidic bond conformations, whereas anti-parallel G-quadruplexes have guanosines that can adopt both *syn*- and *anti*-conformations. The stability of G-quadruplexes is dependent on the presence of monovalent cations, such as potassium and sodium. The association of the four guanine bases creates a central negative electrostatic channel, which is stabilized through the coordination of the metal ion with the *O6* lone pairs of the guanine bases (**Figure 1.5B**).[67] Alternatively, molecular crowding allows the folding of G-quadruplexes under metal-deficient conditions.[68-69] G-quadruplex DNA structures can fold within millisecond under near-physiological conditions,[70] and can exhibit high thermodynamic stability with melting temperatures generally above 50 °C.[71] In addition to its unique structure, G-quadruplexes DNA exhibit characteristic photophysical properties as compared to single-stranded or duplex DNA. These include enhanced quantum yields and photo-excited lifetimes,[72-73] as well as high energy transfer efficiencies.[74-75] These features are the result of potassium ions binding and were revealed by an internal fluorescent probe.[74]

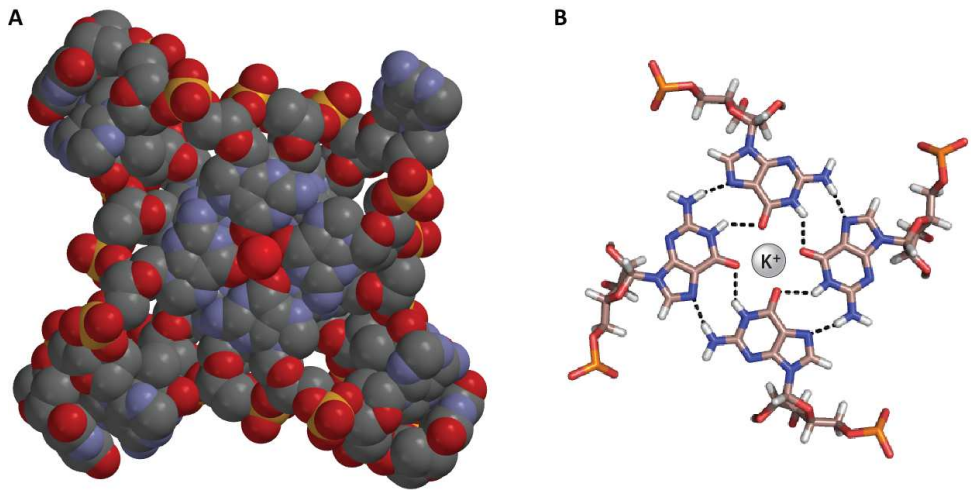


Figure 1.5. Structural representation of G-quadruplex DNA. (A) 3D structure of G-quadruplex DNA as a space filling representation. (B) Tube representation of the G-tetrad base pairing. The structure was obtained from PDB and viewed with Spartan 10.

1.4.2. Evidence for G-quadruplex *in vivo*

A large number of putative G-quadruplex forming motifs $(G_3N_{1-7})_4$ (where G = guanosine and N = any base) can be found throughout the human genome.[76] DNA sequences with the ability to fold into G-quadruplex structures have been implicated in a variety of genome function, which include gene transcription,[77] recombination,[78] chromosome stability,[79] and programmed cell death.[80] The high thermodynamic stability under near-physiological conditions associated with G-quadruplexes structures suggests their potential formation in cells. Indeed, G-quadruplex structures have been identified using antibodies in the nuclei of human cancer cells,[15] and small molecules have been shown to bind to G-quadruplex and inhibit transcriptional activity of the *c-MYC* promoter region.[81-82]

1.5. Metallo DNA Duplex with T-Hg^{II}-T Base Pairs

Another example of nucleobase-metal coordination is provided by DNA duplexes containing mismatched base pairs that selectively bind metal ions. The interaction of Hg^{II} ions with thymidine residues in duplex DNA has been the focus of investigation for many years.[83-84] The binding mechanism of mercury(II) ions (Hg^{II}) with DNA was initially proposed through strand slippage and coordination of Hg^{II} with the *N*3 position of thymidine to form T–Hg^{II}–T base pairs.[83-84] NMR spectroscopy suggested

that Hg^{II} ions interact with A-T base pairs *via* coordination by NH_2 of adenosine and *O4* of thymidine.[85] Recently, Hg^{II} ions have been shown to specifically bind to thymidine mismatches in duplex DNA *via* the formation of T–Hg^{II}–T base pairs.[86] Among many metals known to interact with nucleic acids, only Hg^{II} was able to strongly interact with thymidine mismatches,[86] at a molar ratio of 1:1 and with a reported binding constant of nearly 10^6 M^{-1}.[87] Direct evidence from ^{15}N NMR labeling studies proved that the Hg^{II} ion coordinates to the *N3* position of thymidine following its deprotonation (**Figure 1.6C**).[88] The formation of T–Hg^{II}–T base pairs results in the stabilization of duplex DNA with thermal temperatures comparable to the corresponding Watson-Crick base pair.[89] The crystal and solution structure of the metallo-DNA complex were recently obtained (**Figure 1.6A** and **B**). This has shown that the T–Hg^{II}–T base pairs are very well accommodated in duplex DNA and mimic Watson-Crick base pairs.[90-91] The dehydration of Hg^{II} ion inside the DNA structure is required for formation of the thymidine base pair. The local helical parameters and intra-base pair parameters indicated that the DNA adopted a B-DNA conformation. The *N3*–Hg^{II}–*N3* bond is linear and the distance between the *N3* of thymidine and the Hg^{II} ion is 2.0 Å, with the metal located in the center of the two thymidines.[90] The propeller twist angles of T–Hg^{II}–T base pair (-22 ° and -20 °) are larger than that of Watson-Crick (-1°), due to the repulsion of the carbonyl groups and the absence of hydrogen bonds. In duplex DNA with two consecutive T–Hg^{II}–T pairs, the Hg^{II} ions are aligned along the helical axis and the Hg^{II}–Hg^{II} distance is 3.3 Å, suggesting metallophilic attraction.[92] Since the T–Hg^{II}–T artificial base pair closely resembles Watson-Crick base pairs, DNA polymerase could misincorporate a thymidine triphosphate into the opposing site of a thymidine in the presence of Hg^{II} ions.[93]

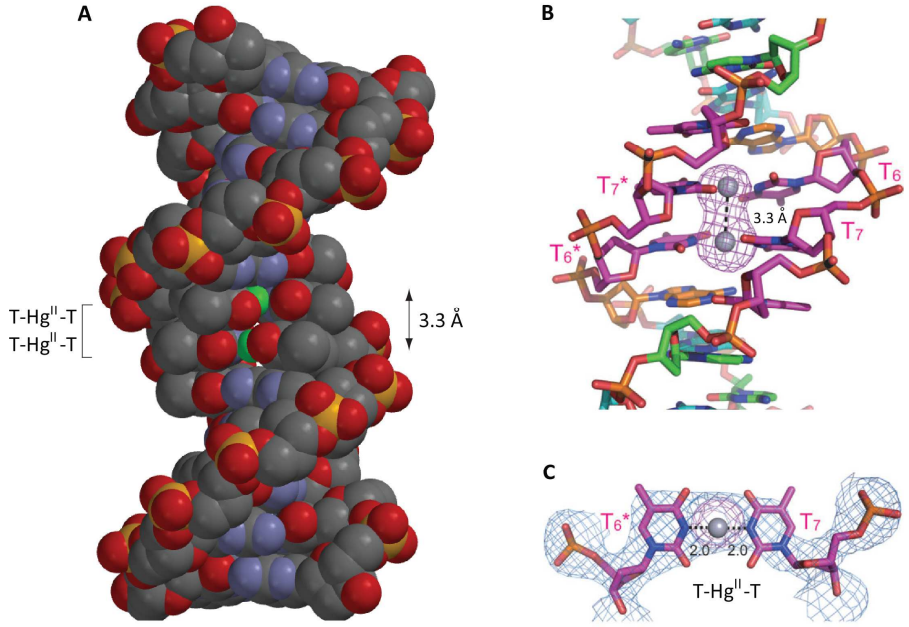


Figure 1.6. Structural representation of metallo DNA duplex with T-Hg^{II}-T base pairs. (A) 3D structure shown as a space filling model obtained from crystal structure as PDB 4L24.[90] Hg^{II} ions are shown in green spheres in this model (B) Side view a local omit map for Hg^{II} ions with Hg^{II}-Hg^{II} ions atom distances. (C) Top view of a T-Hg^{II}-T base pair and atom distances between *N*3 of T and Hg^{II}. Hg^{II} ions are shown in grey spheres in this model. Figures B and C were reproduced from Kondo *et al.*[90]

Positively-charged ions can promote biologically important reactions within DNA structures.[94] Proton, potassium and mercury ions can specifically bind to certain DNA sequences to induce folding into biologically relevant structures. Upon binding to nucleic acids, ions alter the electronic structures of each nucleobase. Fluorescent probes can therefore be incorporated into DNA to report its folding using these characteristic photophysical changes upon ion binding.

1.6. Detection of Nucleic Acid Conformation and Base Pairing

1.6.1. Traditional Biophysical Methods

A wide variety of biophysical techniques has been used to characterize nucleic acids structure and function *in vitro*. Circular dichroism (CD) provides an accessible tool to evaluate DNA secondary structures.[95] CD spectroscopy relies on the chirality of DNA which absorbs differently right- and left-

handed circularly polarized ultraviolet light (λ_{abs} = 180 – 300 nm) to give a CD signal, reported as molar ellipticity (θ). This technique is more sensitive than NMR and can be used with samples diluted to 20 μg/mL. Thermal difference spectra (TDS) provide another simple method for analyzing the secondary structure of DNA.[96] This tool is based on the difference of ultraviolet absorbance spectra of the folded versus unfolded states at temperatures below and above the melting temperature (T_m), respectively. The main drawback of these two techniques is they provide only global folding aspects of DNA structure. In order to obtain highly detailed information at the base pairing level, NMR spectroscopy and X-ray crystallography have been used to characterize nucleic acids structure.[97-98] DNA topologies, base-pair and base step parameters, as well as base-stacking distances can be resolved using these techniques, but the need for large amounts of DNA and high quality crystals have limited their utilization. In addition, most of these methods require pure samples and are incompatible with conformational analyses in living cells.

1.6.2. Fluorescence Spectroscopy

Fluorescence spectroscopy provides the most powerful tool for studying DNA folding, function, and localization in living cells and allows real-time observations of changes in microenvironments.[99-105] It benefits from excellent sensitivity, and a large number of experimental approaches, including steady state and time-resolved intensity measurements, Förster resonance energy transfer (FRET) and fluorescence anisotropy. FRET represents one of the most common techniques to study DNA global conformation changes using fluorescence,[106-107] and fluorescence anisotropy is a powerful tool to measure the dynamic motions of fluorophores.[108] Fluorescent probes that resemble biomolecular building blocks can enable powerful approaches for probing the folding and activities of biological macromolecules.[109] The absence of endogenous fluorescent emissions from DNA molecules offers the advantage that fluorescent signals from unnatural nucleotide base analogs can be observed without competing background signals.[110-111] The main advantages of the use of internal probe include minimal structural perturbations and high resolution information obtained at the level of single nucleotides. Such fluorescent nucleosides have already greatly contributed to the understanding of nucleic acid structure,[112] recognition[113-116] and function.[117] All of these fluorescence-based methods can be applied on single molecules and in whole biological systems to provide information about the environment, dynamics, folding, location and binding of nucleic acids. The three main types of fluorescent probe can be presented as (A) external fluorescent probes, (B) covalent fluorescent probes, and (C) internal fluorescent probes (**Figure 1.7**).

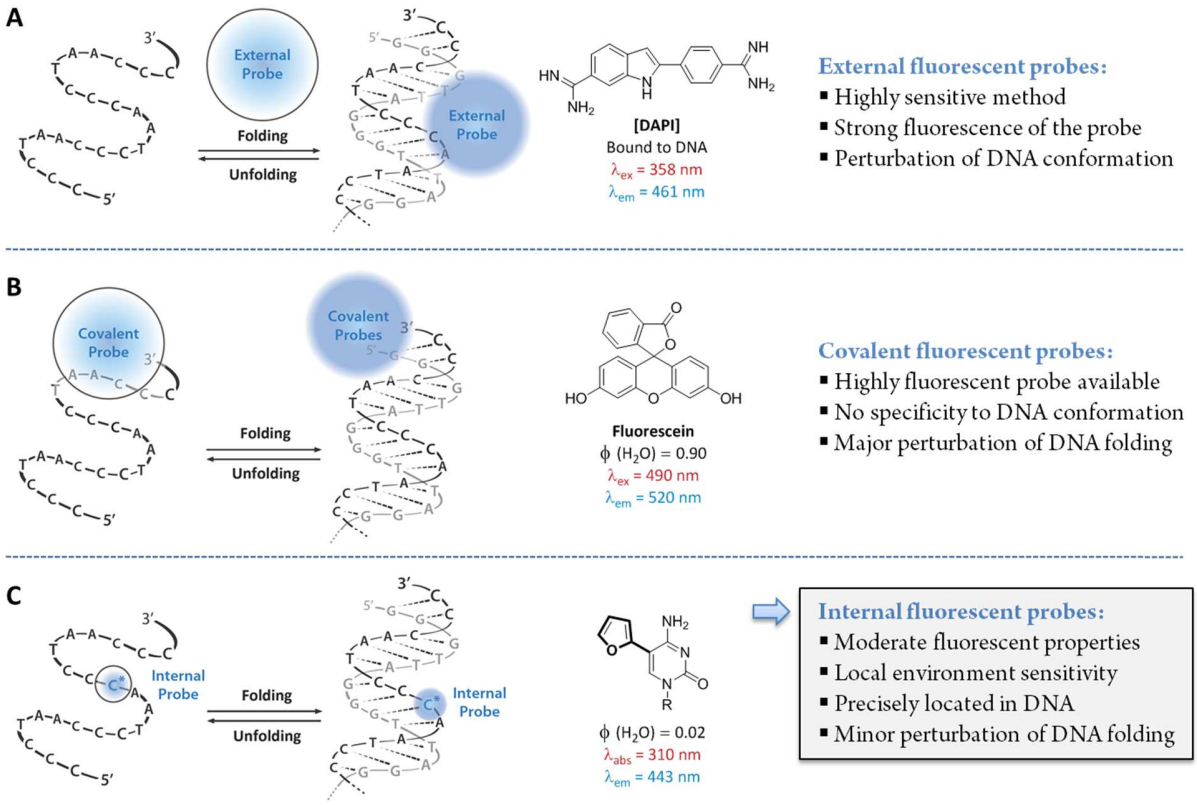


Figure 1.7. Fluorescence labeling strategies illustrated for the detection of duplex DNA. (A) External fluorescent probes, (B) covalent fluorescent probes and (C) internal fluorescent probes.

1.6.3. External Probes

External fluorescent probes which bind to DNA with high affinity through stacking and/or electrostatic interactions can be used to specifically recognize certain structures. The blue fluorescent probe DAPI is a commonly-used agent in fluorescent microscopy that stains preferentially duplex DNAs by binding to the groove regions (**Figure 1.8**). A fluorescence enhancement of ca. 20 fold is observed when bound to DNA (ϕ = 0.34)[118] as compared to the free dye (ϕ = 0.02) through loss of its rotatory motions upon DNA binding. Acridine orange represents another cell-permeant staining agent for nucleic acids.[119] It emits green fluorescence (λ_{em} = 525 nm) when bound to duplex DNA and red fluorescence (λ_{em} = 650 nm) when bound to single-stranded DNA (**Figure 1.8**).[120] At low pH, acridine orange will emit orange light when excited by blue light. The main drawback of these dyes is in the lack of specificity towards DNA conformations. Structure-specific fluorescent probes that bind with high affinity and selectively to G-quadruplex with concomitant increases of fluorescence intensity were developed.[121-122]

Although very sensitive and used *in vivo*, high affinity external probes can perturb the conformation preference of DNA itself and therefore lead to a false readout of the folding.[123-124]

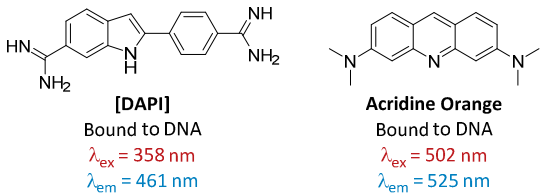


Figure 1.8. Examples of "external" fluorescent dyes for staining cellular DNA.

1.6.4. Covalent Probes

The covalent attachment of large fluorophores to the ends of oligonucleotides represents one of the most widely-used techniques to study DNA folding and dynamics *in vitro*. The technique is based on FRET, which is the transfer of energy from one probe (donor) to the other (acceptor) when they are in close proximity (usually 10 – 100 Å). Fluorescein and rhodamine are very good probes for FRET (**Figure 1.9**) due to their high extinction coefficients ($\varepsilon = 9.9 \times 10^4$ and 6.5×10^4 M^{-1} cm^{-1}, respectively), quantum yields (ϕ = 0.90 and 0.70, respectively) and good spectral overlap (λ_{em} = 520 nm and λ_{abs} = 556 nm, respectively). The main advantage of this technique is sensitivity which comes from the highly desirable photophysical properties of the probes. However, the large modifications made to the DNA by these fluorophores can affect the folding and the thermal stability of the DNA.[124-125] More importantly, the fluorescent signal is limited to the large changes in probe-to-probe distances and is therefore only able to report global aspects of DNA folding.[50]

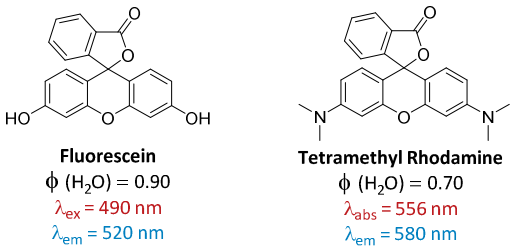


Figure 1.9. Examples of covalently-attached fluorescent probes for FRET experiments in DNA.

1.6.5. Fluorescent Nucleoside Analogs as Internal Probes

As the core heterocycles commonly found in nucleic acids are practically non-emissive,[110] numerous efforts to impart useful photophysical features upon pyrimidines and purines have been reported.[109,126] These aromatic heterocycles are compatible with diverse modifications and were therefore a fertile ground for synthetic organic chemists. In some cases, minimal structural modifications can dramatically alter their photophysical characteristics. A number of criteria must be fulfilled when developing new internal fluorescent probes. The following characteristics represent the desired attributes of the designed nucleosides: (i) A structural similarity to the native nucleobases to mimic their shape, as well as hybridization and recognition properties including the retention of Watson-Crick hydrogen bonding face. (ii) A red-shifted absorption spectrum to minimize overlap with the absorption of the natural bases and for selective excitation in cellular systems. (iii) A red-shifted emission band preferably in the visible range. (iv) Emission solvatochromism induced upon changes in the environment polarity. (v) High quantum efficiency in the context of folded nucleic acids. (vi) Fluorescence sensitivity according to the folded DNA structure. The various families of fluorescent pyrimidine nucleoside analogs can be broadly classified into the following categories:

(A) Fluorescent pyrimidine nucleosides containing expanded nucleobases:[113,127-128] Expanded conjugation of the natural nucleobases by fusing additional aromatic rings onto the pyrimidine nuclei (nucleosides **1** – **4**, **Figure 1.10**). The best examples retain their Watson-Crick hydrogen bonding face, although their large surface area could structurally perturb the resulting oligonucleotides. Having an expanded aromatic surface typically results in favorable photophysical properties, with red-shifted absorption band compared to the natural pyrimidine, emission bands near or in the visible range, and sometimes high emission quantum efficiencies.

(B) Fluorescent pyrimidine nucleosides containing isomorphic nucleobases:[117,129-130] Isomorphic nucleobases analogs are heterocycles that closely resemble to the natural nucleobases with respect to their overall dimensions, hydrogen bonding patterns, and ability to form isostructural Watson-Crick base pairs (nucleosides **5** – **8**, **Figure 1.10**). A clear advantage of these analogs is their strong similarity to the native nucleosides and minimally perturbing nature. However, their photophysical properties are usually less favorable than the expanded nucleobases with lower quantum yields and extinction coefficients.

(C) Pyrimidine nucleosides containing conjugated fluorophores:[131-132] Connecting known chromophores *via* non-conjugated linkers yields nucleoside analogs with photophysical features that are normally very similar to that of the parent fluorophore. The use of conjugated linkers typically generates a new chromophore with unique photophysical characteristics. The main drawback of this class of nucleobase is the poor resemblance to the parent nucleosides which can lead to structural perturbations upon its incorporation into DNA.

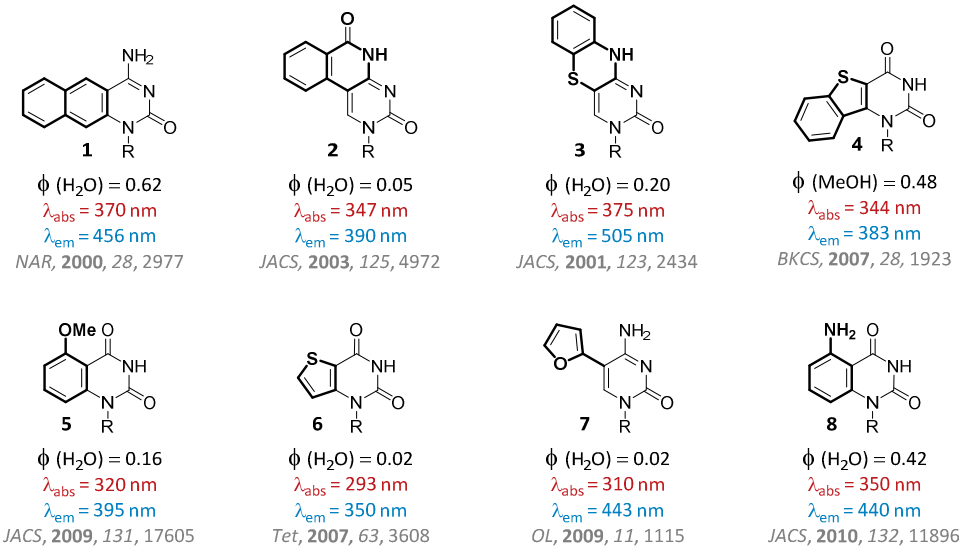
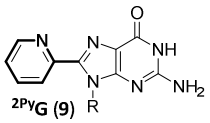


Figure 1.10. Selected examples of fluorescent pyrimidine nucleosides.

Examination of the photophysical characteristics of pyrimidine analogs reveals a common and expected trend. High emission quantum efficiencies and long emission wavelengthes are typically associated with significant structural deviations as compared to the natural uracil or cytosine nucleobases (i.e. compound **1, Figure 1.10**). In contrast, when close structural homology is maintained, the photophysical characteristics are usually less favorable (i.e. compound **6, Figure 1.10**). The design of new fluorescent base analogs therefore requires a delicate balance of minimizing structural changes while maximizing quantum yields, wavelengths of excitation/emission, and environmental sensitivity. Despite this diversity in structures and photophysical properties, the implementation of fluorescent pyrimidine nucleoside analogues in biophysical assays remains challenging, as there is no single analogue that can be universally employed.

1.6.6. Detection of G-quadruplex Folding

In our laboratory, the development of a sensitive method for the detection of G-quadruplex folding has been achieved using the highly fluorescent guanosine mimic ^{2Py}**G** (**9**).[133] This compound differs from the natural nucleobase by a single pyridine unit attached to the *C8* position of guanine (**Figure 1.11**). The ^{2Py}**G** nucleoside exhibits environment-sensitive photophysical characteristics including an excellent correlation between Stokes shift and solvent polarity. These suggests the presence of solvent-mediated stabilization of a charge separated emissive state that has a larger dipole moment than the ground state. The quantum yield of ^{2Py}**G** strongly correlates to solvent acidity, consistent with an excited state proton transfer as a dominant non-radiative decay pathway of ^{2Py}**G** in aqueous and organic solvents. Because of the large changes in hydration that can occur during conformational changes in DNA/RNA molecules, ^{2Py}**G** was able to serve as a conformation-specific fluorescent probe to detect G-quadruplex DNA folding. ^{2Py}**G** was found to be minimally disruptive to both duplex (ΔT_m = -3.5 °C on average) and G-quadruplex structures (ΔT_m = -1.0 °C on average) as determined by temperature-dependent circular dichroism measurements. ^{2Py}**G** was the first reported example of a fluorescent guanine mimic that exhibits a higher quantum yield upon incorporation into folded oligonucleotides (ϕ = 0.03 – 0.15) as compared to the free nucleoside in water (ϕ = 0.02). Due to good spectral overlap, ^{2Py}**G** ($\lambda_{ex} \approx$ 330 nm in DNA) can act as an energy acceptor for unmodified guanine residues ($\lambda_{em} \approx$ 330 nm in DNA). This internal probe revealed the remarkable energy transfer properties of G-quadruplexes DNA (η_t = 0.11 - 0.41) as compared to duplex DNA (η_t = 0.01 – 0.07).[74]



^{2Py}**G** (**9**)

ϕ (H_2O) = 0.02

λ_{abs} = 305 nm

λ_{em} = 415 nm

JACS, **2010**, *132*, 18004

Figure 1.11. Structure and photophysical properties of the fluorescent guanosine mimic ^{2Py}**G**.

1.7. Objectives: Detection of *i*-motif Folding

Today, the potential biological relevance of the equilibrium between duplex DNA and G-quadruplex / *i*-motif structures remains an open question. In order to gain insights into this equilibrium, we decided to develop a sensitive method that can accurately report *i*-motif folding. Our approach focused on the incorporation of a fluorescent cytosine mimic into *i*-motif DNAs at positions directly involved in C-C$^+$ base pairs. The successful implementation of a fluorescent probe for *i*-motif folding will shed light on the fundamental properties of this DNA structure, and its potential function in cells.

The design and incorporation of cytidine nucleoside analogs that can serve as a selective "turn on" fluorescent probes for folded *i*-motif DNA structures are required to achieve these goals. In this work, the combination of modern organic chemistry with fluorescence spectroscopy were used to address important questions about *i*-motif DNAs. The thesis is therefore divided into the following chapters:

I. An introduction to DNA conformations and biological functions is presented in **Chapter 1**. The detailed analysis of non-duplex structures, such as *i*-motifs and G-quadruplex is emphasized. The traditional and emerging methods for the detection of nucleic acids are reported.

II. The design and synthesis of minimally-disruptive pyrimidine nucleosides with high structural resemblance to the native nucleoside are presented in **Chapter 2**. To access the designed nucleoside analogs in high yields, the development of practical and scalable synthetic routes to 2'-deoxynucleosides were developed. The results described in this chapter were published as part of: Mata, G. and Luedtke, N. W. *J. Org. Chem.* **2012**, *77*, 9006-9017.

III. In **Chapter 3**, the evaluation of the photophysical properties and microenvironment sensitivities of biaryl nucleosides are presented. This chapter describes the relationships between the structure and the photophysical properties of each thymidine analog to identify the most promising emissive candidate. The results described were published as part of: Mata, G. and Luedtke, N. W. *Org. Lett.* **2013**, *15*, 2462-2465.

IV. The successful implementation of a fluorescent cytosine mimic (^{DMA}**C**) into the telomeric cytosine rich DNA sequence that folds into an *i*-motif structure is reported. The evaluation of the impact of ^{DMA}**C** in the context of this highly demanding proton-folded structure is presented in **Chapter 4**. This implementation resulted in the development of a fluorescence-based assay for the detection of static folded DNA and for the real-time study of conformational dynamics between *i*-motif and duplex structures. The results described in this chapter were published as part of: Mata, G. and Luedtke, N. W. *J. Am. Chem. Soc.* **2015**, *137*, 699-707.

V. In **Chapter 5**, additional fluorescence characteristics of ^{DMA}**C** in duplex DNA are presented. The probe was used to detect base pairing mismatches for typing of single nucleotide polymorphisms (SNPs).

VI. The implementation of a fluorescent thymidine mimic (^{DMA}**T**) into DNA structure is presented in **Chapter 6**. In addition to reporting static structure, ^{DMA}**T** was used as a sensitive fluorescent tool for the detection of Hg^{II}-mediated base pair within duplex DNA.

VII. **Chapter 7** combined the materials and methods, computational details, synthetic procedures, characterization of isolated compounds, fluorescence, circular dichroism (CD) and thermal denaturation (T_m) experimental details. A collection of additional CD, T_m, and fluorescence spectra is provided in the Appendixes section.

Chapter 2 | Stereoselective Synthesis of Nucleosides *via N*-Glycosylation of Thioribosides

Chapter 2

Stereoselective Synthesis of Nucleosides *via N*-Glycosylation of Thioribosides

2.1. Summary

An efficient method for the *N*-2-deoxyribosylation of modified nucleobases by 2-deoxy-thioriboside donors is reported (**Figure 2.1**). In the presence of an in situ silylated nucleobase, thioglycosides can be activated with NIS/HOTf to give nucleosides in high yields and with good β-selectivity. By tuning the protecting groups on the C3 and C5 hydroxyls, α/β ratios ranging from 1.0:4.0 to 4.5:1.0 can be obtained. This strategy is applicable to the synthesis of various nucleosides, including ring-expanded pyrimidine derivatives containing sulfur that have previously been reported in low yields. The utility of this approach is further demonstrated by the synthesis of fluorescent nucleosides analogues such as quinazoline and oxophenothiazine that should find broad utility in DNA folding and recognition studies.

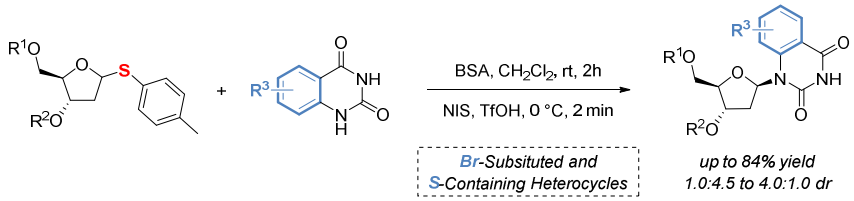


Figure 2.1. Stereoselective *N*-glycosylation of a thioriboside.

2.2. Introduction

Due to their broad spectrum of antiviral and antitumor activities, the synthesis of nucleoside analogues has been a key focus of medicinal chemistry for nearly 40 years.[134-136] More recently, modified nucleosides and nucleotides have emerged as tools for manipulating genetic processes,[137-139] as fluorescent probes for studying DNA folding and recognition,[74,109] and as metabolic labels for cellular DNA *in vivo*.[100,140] In many cases, these applications require synthetic nucleosides with ring-expanded nucleobases that exhibit unique photophysical and/or chemical reactivities.

Synthetic strategies that access 2'-β-deoxynucleosides typically involve *N*-glycosylation of ribofuranosyl derivatives. Stereoselectivity for the desired β-anomer can be achieved by using 1-*O*-acetyl-ribofuranosyl derivatives that mediate a C2-*O*-ester neighboring group effect (**Figure 2.2**, **route A**). Using this approach, high stereoselectivity and good yields can be obtained, but subsequent 2'-deoxygenation procedures must be performed to obtain the desired 2'-D-deoxyribonucleosides. This requires 2-4 additional synthetic steps therefore resulting in only low to moderate overall yields.[141-144] In

addition, commonly used deoxygenation reactions that proceed via radical elimination are not compatible with haloheterocyclic derivatives.[145-146] An alternative approach for 2'-deoxygenation using a photosensitized electron-transfer reaction has recently been reported,[147-148] but this elegant strategy is incompatible with fluorescent nucleobases that act as electron acceptors.[149]

Glycosylation reactions that employ protected 2-deoxyribosides as glycosyl donors provide a more direct pathway to the synthesis of 2'-β-deoxynucleosides, but previously reported examples of *N*-2-deoxyribosylation suffer from modest yields and poor β-selectivity. The most commonly reported approach utilizes the commercially available 2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranosyl chloride as the glycosyl donor (**Figure 2.2**, **route B**). Although it has been successfully used for the stereoselective glycosylation of purine analogues via an SN_2-type mechanism under basic conditions,[150-151] this donor suffers from a number of significant drawbacks for pyrimidine derivatives under Lewis acidic conditions.[152-153] The long reaction times needed for reacting with weak nucleophiles can result in epimerization and the formation of α-nucleosides as the major product.[154-155] Its instability in solution can also result in low to moderate yields of the isolated coupling products, as well as inseparable α/β anomeric mixtures of nucleosides.[155-160] Moreover, direct modification of the C3 and C5 hydroxyl groups is not normally feasible due to the lability of the anomeric chloride. The development of new and efficient stereoselective reactions for constructing 2'-β-D-deoxyribonucleosidic linkages therefore remains an important goal.

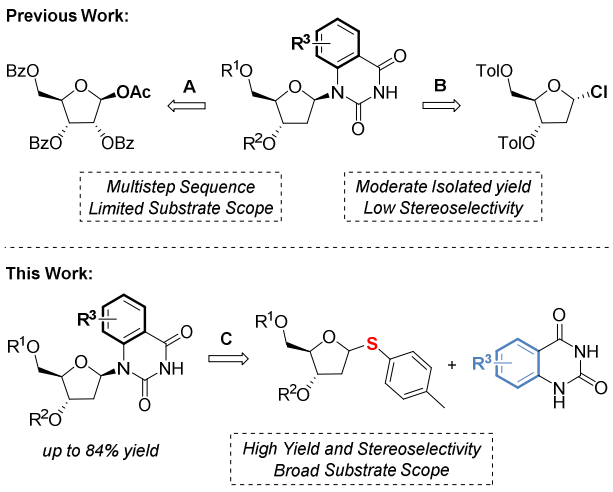


Figure 2.2. Strategies for stereoselective *N*-glycosylation of heterocyclic acceptors.

In recent decades, extensive efforts in oligosaccharide chemistry have revealed a wide variety of efficient glycosylation reactions for the stereoselective synthesis of *O*-glycosidic bonds in complex carbohydrates (**Figure 2.3**).[161] These include the utilization of classic glycosyl donors such as glycosyl bromides, chlorides, acetates, trichloroacetimidates, thioglycosides, and, more recently, the introduction of glycosyl sulfoxides, iodides and phosphites.[162-166] Relatively few of these established methodologies have been applied for *N*-glycosylation reactions or in nucleoside synthesis. A small handful of 2-deoxyribose thioglycosides have been developed as glycosyl donors, but have been used only for the *N*-glycosylation of natural pyrimidine nucleobases which are known to exhibit facile reactions with a wide variety of glycosyl donors.[167-169] In contrast, unnatural nucleobase acceptors with extended surface areas are highly problematic *N*-glycosidic acceptors.[170]

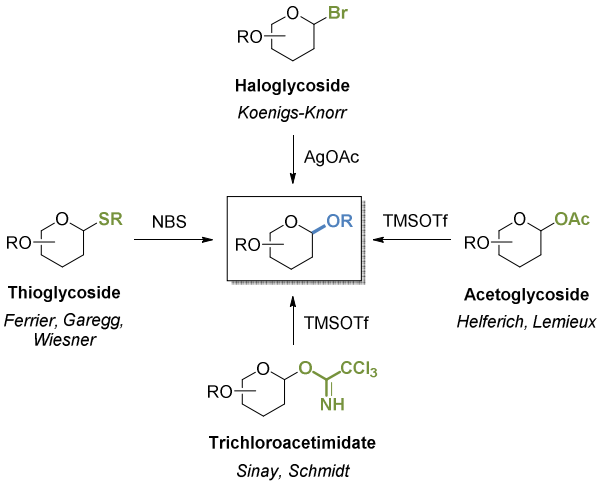


Figure 2.3. Common glycosyl donors and activation conditions for *O*-glycosylation in pyranose series.

Given the growing importance of fluorescent 2'-deoxynucleosides,[109] we become interested in the development of a new 2-deoxyribose donor for robust β-selective *N*-glycosylation reactions in high yields. Here we report Vorbrüggen-type[171-173] *N*-2-deoxyribosilation of 2-deoxy-thioribosides by ring-expanded nucleobase analogues that exhibit limited solubility and poor nucleophilicity (**Figure 2.2**, **route C**). 6-Bromo-quinazoline-2,4-(1*H*,3*H*)-dione was selected as an initial model for evaluating these *N*-glycosylation reactions since quinazoline-2,4-(3*H*)-dione β-nucleosides are synthesized in low yield,[155] and are of particular interest for further elaboration into fluorescent nucleoside analogs. Following optimization of the thioglycoside donor, a variety of other, highly challenging, nucleobases were glycosylated in moderate to high yields and good β-selectivity.

2.3. Results and Discussion

2.3.1. Synthesis of 2-Deoxy-Thioriboside Donors

Our study started with the preparation of variable 2-deoxy-thioriboside donors by systematic variation of the functional groups at the C3 and C5 positions. Thioglycosides are especially amenable to this approach because of their high stability under a wide range of reaction conditions. To generate the thioether aglycon, commercially available 1,3,5-tri-*O*-acetyl-2-deoxy-D-ribose **10** was treated with 1.02 equiv of *p*-toluenethiol (*p*TolSH) and $BF_3{\cdot}Et_2O$ at -78 °C to afford thioglycoside donor **11** (**Figure 2.4**). The resulting 2-deoxy-thioriboside was obtained in 92% yield on a 30 g scale as an anomeric mixture (α/β = 1.0:1.8). Deacetylation of **11** was then conducted using K_2CO_3 in a mixture of MeOH/CH_2Cl_2 to give diol **12** in 98% yield. Regioselective silylation of 5-OH with triisopropylsilyl chloride (TIPSCl) afforded compounds **13β** (56%) and **13α** (31%), which were readily separated using preparative scale silica gel chromatography.

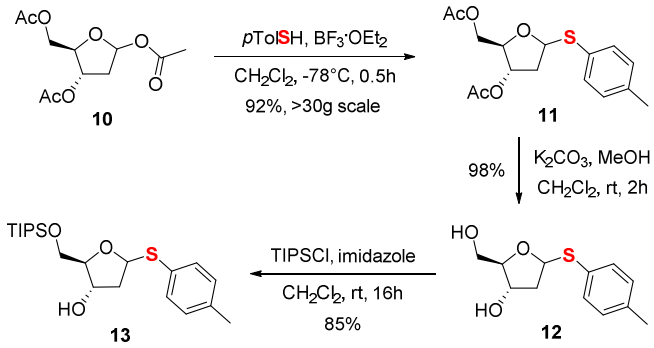


Figure 2.4. Formation of the thioether aglycon.

Based upon previous studies,[174] we anticipated that an ester group at C3 may provide anchimeric assistance, enhancing the β-selectivity of *N*-glycosylation. C3-*O*-esterification reactions were therefore performed with the combined thioglycosides **13β** and **13α** to afford esters **14** – **16** (**Figure 2.5**) that were obtained as anomeric mixtures (α/β = 1.0:1.8). Thioglycoside **14** was obtained in 90% yield by acetylation with Ac_2O, and thioglycosides **15** and **16** were obtained after treatment with benzoyl chloride (BzCl) or *p*-methoxybenzoyl chloride (PMBCl), in 87% and 78% yields, respectively. Based on previous studies by Zhang and co-workers,[175] we were interested in a possible neighboring group participation by an *N*-

acetyl-glycine residue at the C3 position. Enantiopure thioglycosides **17α** and **17β** containing C3-*O*-*N*-acetyl-glycine were therefore prepared using *N*,*N*'-diisopropylcarbodiimide (DCI) as a coupling reagent.

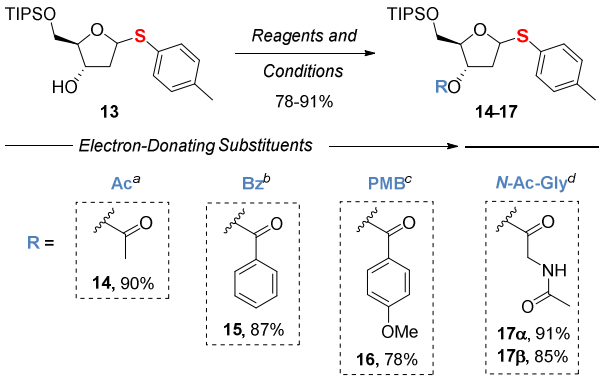


Figure 2.5. Synthesis of C3-substituted 2-deoxy-thioriboside donors **14-17**. Reagents and conditions: a) Ac_2O, Et_3N, DMAP, MeCN, rt, 2h; b) BzCl, Et_3N, DMAP, pyridine, CH_2Cl_2, rt, 2h; c) PMBCl, Et_3N, DMAP, pyridine, CH_2Cl_2, rt, 16h; d) N-Ac-Gly-OH, DCI, DMAP, CH_2Cl_2, rt, 12h.

Based on previous studies,[176] we also focused our attention on C3-5-*O*-silylated cyclic functional groups that can induce conformational restrictions and therefore influence stereoselectivity during glycosylation. C3-5-*O*-silylated thioglycosides **18** and **19** containing six- and eight-membered rings were therefore prepared from **12** in 83% and 93% yield, respectively (**Figure 2.6**). Finally, in order to provide a direct comparison of differences in chemical reactivity between thioglycosides and the commonly used 2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranosyl chloride, compound **20** was synthesized in 86% yield.

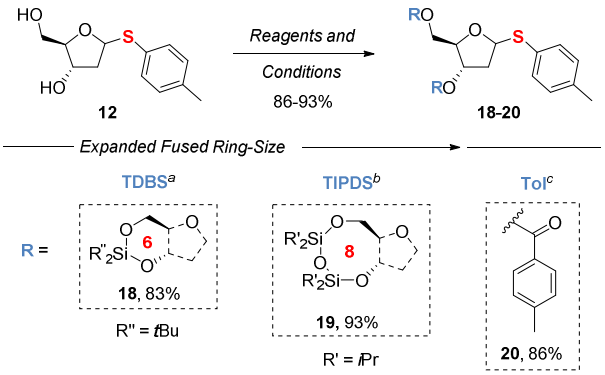


Figure 2.6. Synthesis of C3,5-O-dilylated 6- and 8-membered ring 2-deoxy-thioriboside donors **18** – **19** and **20**. Reagents and conditions: a) tBu_2SiOTf_2, 2,6-lutidine, CH_2Cl_2/DMF, 0 °C to rt, 16h; b) TIPDSCl, pyridine, 0 °C to rt, 12h; c) *p*-toluoyl chloride, Et_3N, DMAP, pyridine, CH_2Cl_2, rt, 16h.

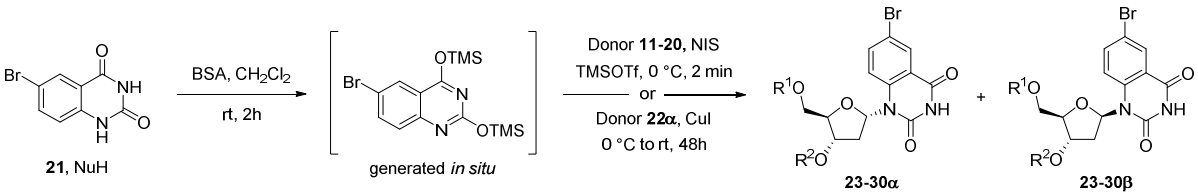
2.3.2. *N*-2-Deoxyribosilation of 2-Deoxy-Thioriboside Donors

With the 2-deoxy-thioriboside donors **11**, and **14** – **20**, in hand, we investigated their reactivity in the *N*-glycosylation of 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione (**21**, NuH). This reaction was carried out starting with *in situ* silylation of the acceptor nucleobase **21** (NuH, 1.2 equiv) with *N*,*O*-bis(trimethylsilyl)acetamide (BSA, 2.5 equiv) in CH_2Cl_2 for 2h, followed by addition of the 2-deoxy-thioriboside donor (**11** – **20**, 1.0 equiv) and sequential addition of *N*-Iodosuccinimide (NIS, 1.2 equiv) and trimethylsilyl trifluoromethanesulfonate (TMSOTf, 0.6 equiv).[177] The isolated yields and α/β ratios (as determined by ^{1}H NMR of the crude reactions) of these reactions are summarized in **Table 2.1**.

When thioglycoside **20** (α/β = 1.0:1.8) was activated by a combination of NIS and TMSOTf, the desired coupling products **23α** and **23β** were isolated in 72% yield with the β-anomer as the major product (α/β = 1.0:1.6). While this diastereoisomeric mixture was not separable using chromatographic methods, the major β-nucleoside **23β** could be isolated via precipitation from EtOAc. Based upon the α/β ratio and total isolated yield, the desired β-nucleoside **23β** was obtained in 43% yield. In comparison, the commonly-used α-glycoside chloride **22α** was activated under standard conditions by CuI in a 48h reaction to afford 58% yield of the expected product, but in a highly α-selective fashion (α/β = 3.5:1.0). This is consistent with previous reports on *N*-glycosylation of quinazoline-2,4-(3H)-dione by **22α**,[155] where the β-nucleoside **23β** was obtained in only 13% yield. Similar poor yields for *N*-glycosylation of other nucleobase acceptors by **22α** have also been reported.[157,160]

The use of thioglycoside **11** (Entry 3) afforded the coupling products **24α** and **24β** in a similar overall yield (76%) as thioglycoside **20** but the β-selectivity was lost (α/β = 1.2:1.0). As compared to **11**, thioglycoside **14** exhibited enhanced β-selectivity (α/β = 1.0:1.8) and a slightly higher isolated yield of 84%. The introduction of a C5-*O*-TIPS group into thioglycoside **14** also enabled facile and quantitative separation of the α/β anomeric mixture from multi-gram reactions using silica gel chromatography (TLC ΔR_f = 0.2) to afford 54% isolated yield of the β-nucleoside **25β**, after only a 2 minute reaction time. The relative regio- and stereochemistry of **25α** and **25β** was confirmed by 2D ^{1}H-^{13}C HMBC and ^{1}H-^{1}H ROESY experiments (**Figure 2.7**, **Figure 2.8**, **Figure 2.9** and **Figure 2.10**). Consistent with an SN_1 pathway involving an oxocarbenium ion intermediate, enantiopure thioglycoside **14β** was submitted to

the same reaction conditions and afforded the exact same anomeric ratio of nucleoside as thioglycoside **14** (α/β = 1.0:1.8).



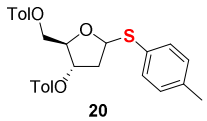
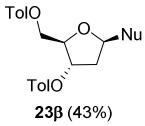
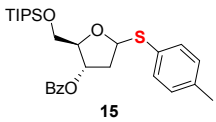
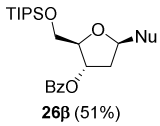
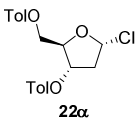
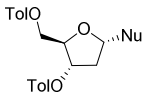
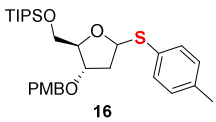
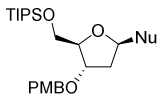
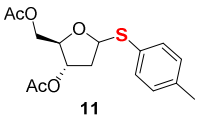
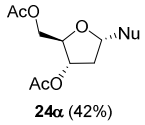
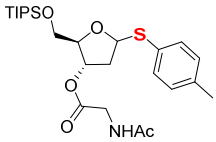
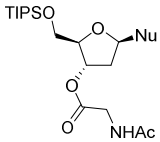
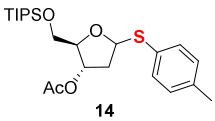
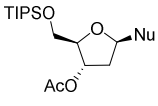
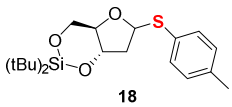
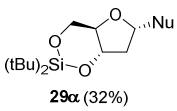
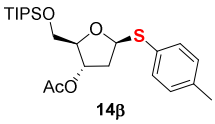
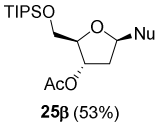
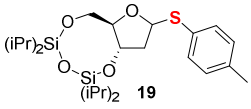
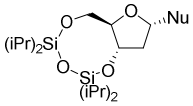
Entry	Donor	Major Product	Yield[a]	α/β[b]	Entry	Donor	Major Product	Yield[a]	α/β[b]
1	 **20**	 **23β** (43%)	72%	1.0:1.6	6	 **15**	 **26β** (51%)	83%	1.0:1.6
2	 **22α**	 **23α** (45%)	58%	3.5:1.0	7	 **16**	 **27β** (50%)	80%	1.0:1.7
3	 **11**	 **24α** (42%)	76%	1.2:1.0	8	 **17α** or **17β**	 **28β** (41%)	51%	1.0:4.0
4	 **14**	 **25β** (54%)	84%	1.0:1.8	9	 **18**	 **29α** (32%)	50%	1.8:1.0
5	 **14β**	 **25β** (53%)	82%	1.0:1.8	10	 **19**	 **30α** (56%)	68%	4.5:1.0

Table 2.1. *N*-2-deoxyribosilation of 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione (**21**, NuH) by 2-deoxy-thioriboside donors. [a] Isolated yield for both anomers. [b] Ratio determined by ^{1}H-NMR analysis of the crude mixture.

The modest differences in yield and stereoselectivity exhibited by donor **14** versus **11** can be explained by the armed-disarmed concept where the electron-withdrawing C5-*O*-acetyl group present in donor **11** is replaced by the electron-donating C5-*O*-TIPS group.[178] This change can "arm" the glycosyl donor by stabilizing the formation of the oxocarbenium ion intermediate (**Figure 2.11**).

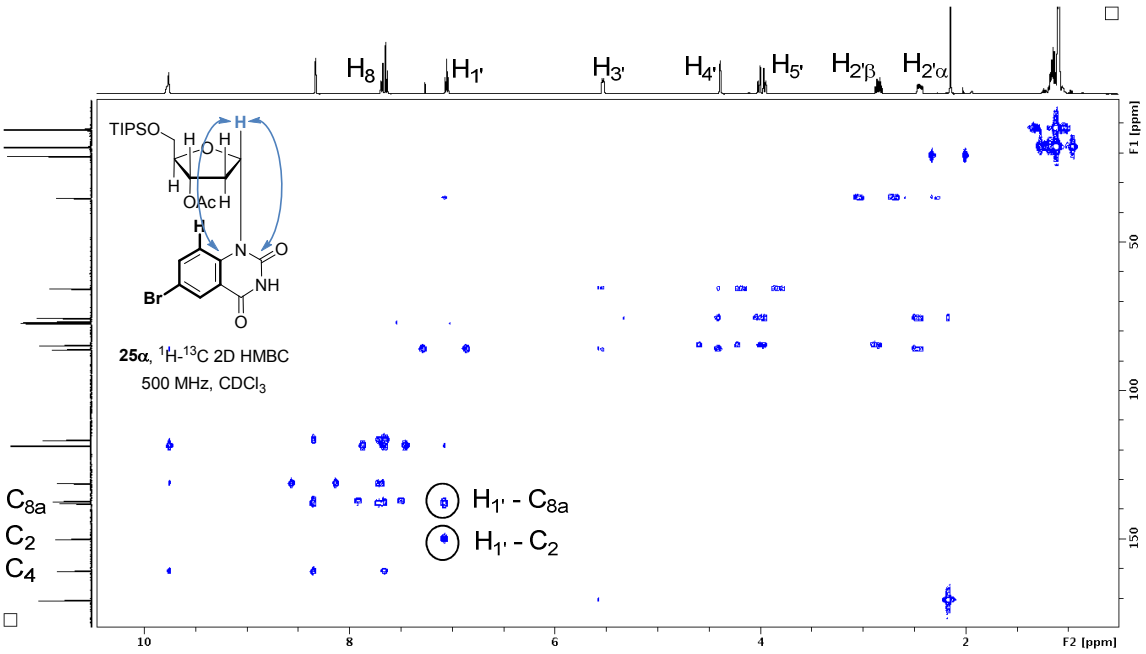


Figure 2.7. Determination of connectivity at *N1* according to the 2D ^{1}H-^{13}C HMBC spectra of **25α**.

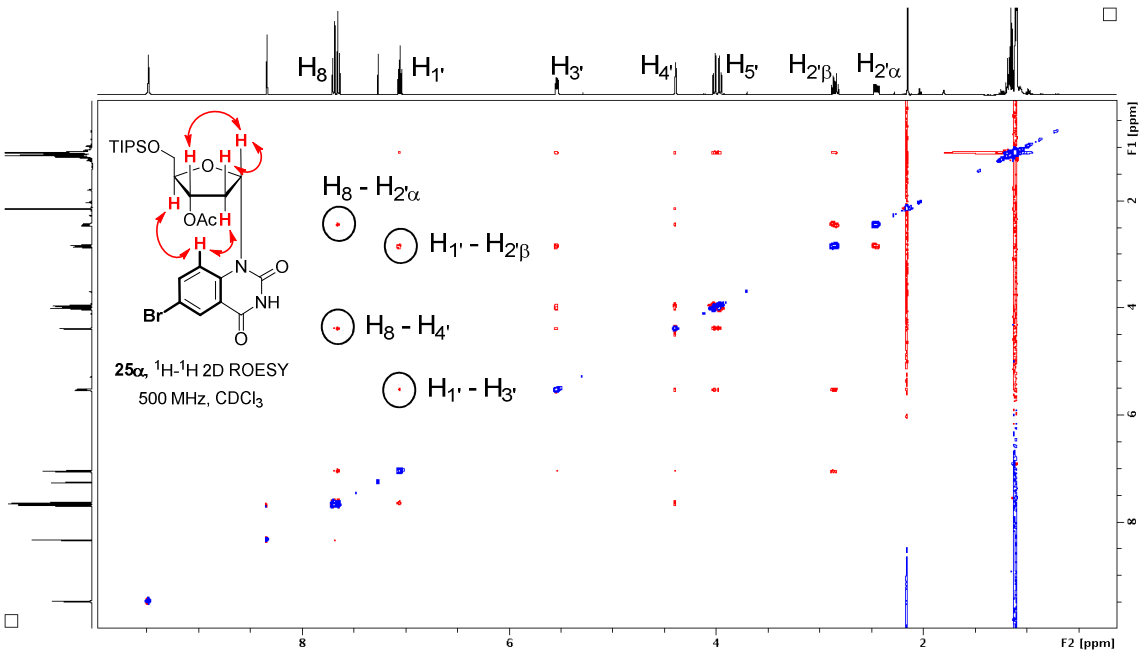


Figure 2.8. Determination of α-anomeric configuration with the 2D ^{1}H-^{1}H ROESY spectra of **25α**.

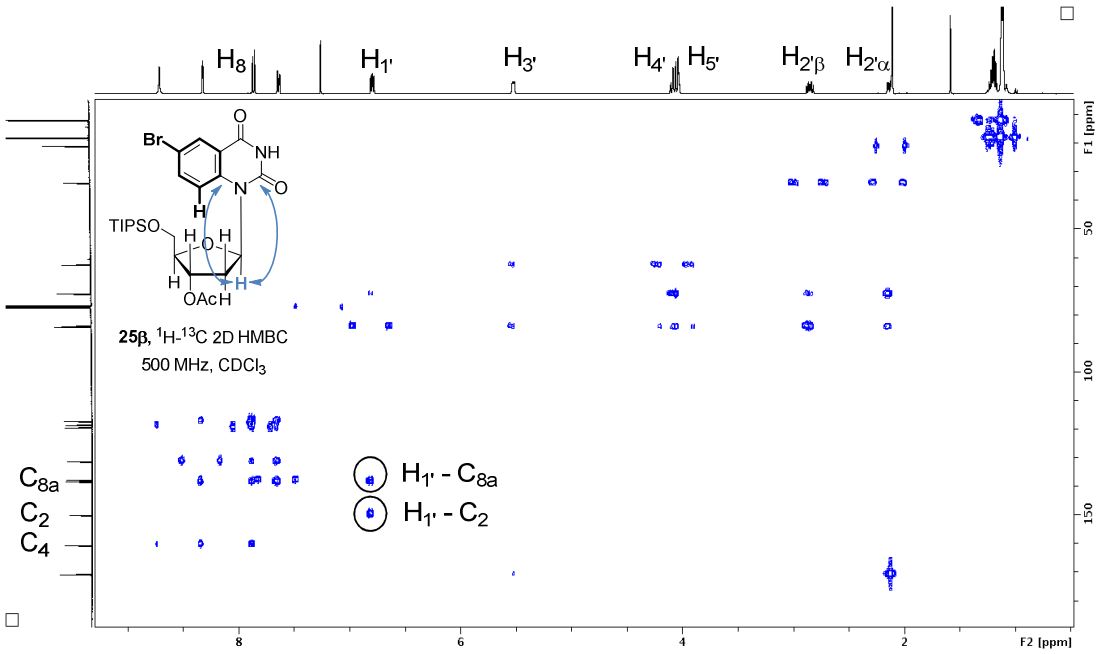


Figure 2.9. Determination of connectivity at *N1* according to the 2D ^{1}H-^{13}C HMBC spectra of **25β**.

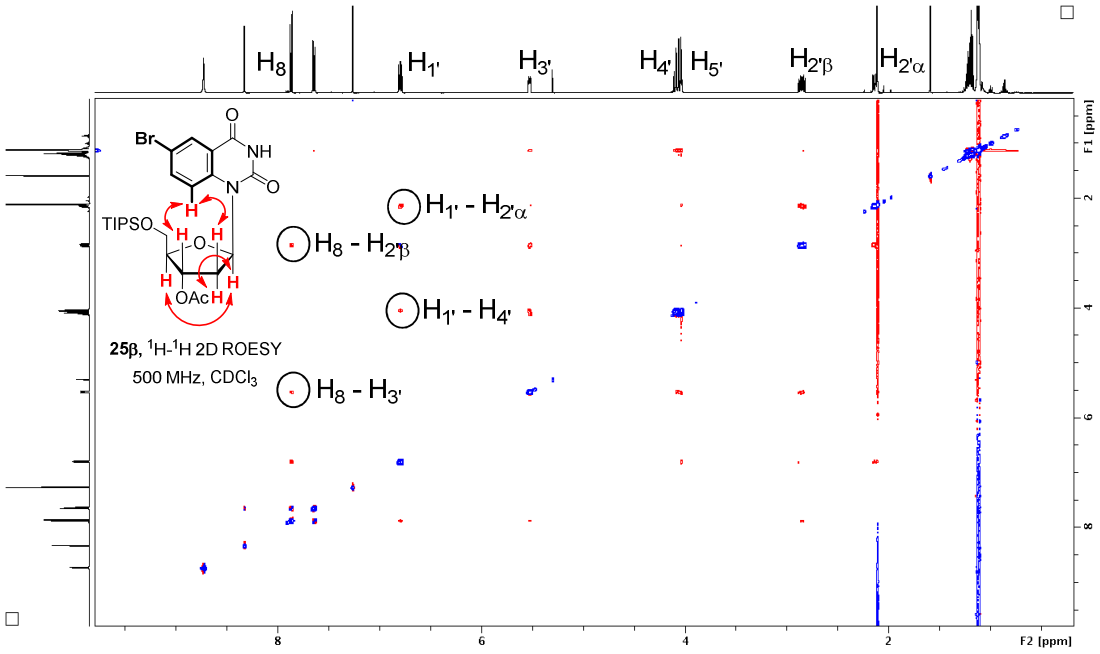


Figure 2.10. Determination of β-anomeric configuration with the 2D ^{1}H-^{1}H ROESY spectra of **25β.**

To further evaluate the armed-disarmed concept in the context of 2'-deoxyribonucleoside synthesis, a competition experiment was conducted where a mixture of thioglycosides **11** (1 equiv) and **14** (1 equiv) competed in the same reaction mixture with a limited amount of the nucleobase (**21**, 1 equiv) and using an excess of activator (NIS, 3 equiv). Consistent with faster oxocarbenium ion formation by thioglycoside **14**, a two-fold difference in product ratio was observed in the crude product mixture (**24**/**25** = 1.0:2.0).

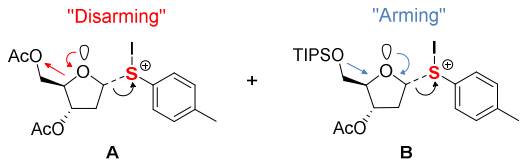


Figure 2.11. Arming and disarming effect in thioglycosides **11** and**14**.

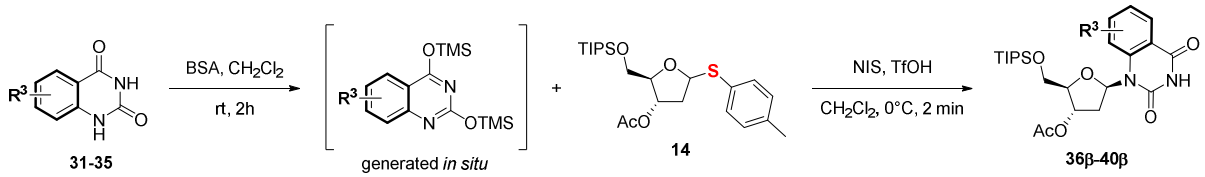
Use of thioglycosides **15** and **16** furnished the expected products in high yields, 83% and 80%, respectively (Entry 6 and 7, **Table 2.1**). Contrary to our expectations,[174] a constant β-selectivity was observed for the benzoyl ester as well as for the more electron-donating *p*-methoxybenzoyl ester at C3 position. These results suggest that simple C3-*O*-ester groups do not act as directing groups in these reactions. However, donors **17α** and **17β** containing *N*-acetyl-glycine groups,[175] exhibited highly stereoselective formation of the β-anomer (α/β = 1.0:4.0), albeit with lower overall isolated yield of 51% under these reaction conditions.

In contrast to **17**, thioglycoside **18** containing the C3-5-*O*-silylated six-membered ring was found to be α-selective (α/β = 1.8:1.0, Entry 9, **Table 2.1**). In this reaction, the coupling products **29α** and **29β** were obtained in 50% yield and unreacted starting material was recovered. More importantly, thioglycoside **19** containing the C3-5-*O*-silylated eight-membered ring gave the corresponding nucleosides **30α** and **30β** in 68% yield in a highly α-selective fashion (α/β = 4.5:1.0).

2.3.3. Scope of the *N*-2-Deoxyribosylation of Ring-Expanded Pyrimidine Nucleobases

Having indentified thioglycoside **14** as an excellent donor for high-yielding, β-selective *N*-2-deoxyribosylation of 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione, we explored the scope of this reaction in constructing 2'-β-D-deoxyribonucleosidic linkages known to be highly challenging. *N*-Heterocyclic nucleobases **32** – **35** were therefore synthesized according to known procedures from their corresponding starting materials **41** – **44**. The corresponding β-nucleosides **36β** – **40β** have been shown to exhibit useful

fluorescent properties upon incorporation into DNA.[179] However, all previous syntheses of these nucleosides have been achieved in low yields by using the α-glycoside chloride method for **38β** (isolated after 3 steps in 9% and 3%[156-157]), the 2'-deoxygenation method for **39β** (isolated after 5 steps in 16%),[143] or the sodium salt method for **40β** (isolated after 2 steps in 14%).[160,170]



Entry	Precursor	Acceptor	Major Product	Yield[a]	α/β[b]	β-Yield
1	/	**31**	**36β**	83%	1.0:1.8	53%
2	**41**	**32** (1 step, 90%)	**37β**	80%	1.0:1.7	50%[c]
3	**42**	**33**[d] (1 step, 82%)	**38β**	70%	1.0:1.8	45%
4	**43**	**34**[d] (1 step, 70%)	**39β**	73%	1.0:2.0	49%[c]
5	**44**	**35**[d] (2 steps, 22%)	**40β**	46%	1.0:2.5	33%

Table 2.2. Scope of the *N*-2-deoxyribosylation of thioriboside **14** with ring-expanded pyrimidine nucleobases. [a] Isolated yield of the combined anomers. [b] Ratio determined by ^{1}H-NMR analysis of the crude mixture. [c] The direct separation of the anomers was not realized. The yield was calculated from the isolated yield of the mixture and the α/β selectivity of the reaction. [d] The nucleobase was first presilylated in HMDS at 110 °C for 10h.

Table 2.2 summarizes our results for *N*-2-deoxyribosylation using thioglycoside donor **14**. Reactions with quinazoline-2,4-(1*H*,3*H*)-dione **31** and 5-methoxyquinazoline-2,4-(1*H*,3*H*)-dione **32** furnished the expected nucleosides in high yield, 83% and 80%, respectively, and with good β-selectivity (**36α**/**36β** = 1.0:1.8 and **37α**/**37β** = 1.0:1.7). Optimal yields were obtained by increasing the solubility of the heterocyclic nucleobase **33** – **35** by presilylation with hexamethyldisilazane (HMDS) at 110°C before being subjected to the standard reaction conditions. Benzo[g]quinazoline-2,4-(1*H*,3*H*)-dione **33**, containing an even larger aromatic surface gave the desired products (**38α**/**38β** = 1.0:1.8) in 70% yield.

Given the use of NIS for thioglycoside activation in these reactions, cross reactivity with sulfur-containing groups was a particular concern. However, thieno[3,2-d]pyrimidine-2,4-dione **34** containing a thiophene group furnished the desired nucleosides in 73% yield and in good β-selectivity (**39α**/**39β** = 1.0:2.0). Furthermore, when thioether **35** was subjected to glycosylation with **14**, the desired nucleosides were obtained in 46% yield with the β-anomer as the major product (**40α**/**40β** = 1.0:2.5). The structure of **40β** was confirmed by 2D ^{1}H-^{1}H ROESY, where correlation between H1' and H4', H1' and H2α', H6 and H1', H6 and H2β' as well as H6 and H4' were observed (**Figure 2.12**).

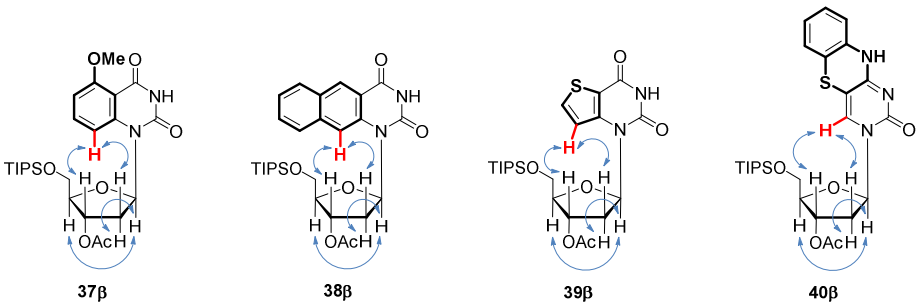


Figure 2.12. Stereochemical proofs with 2D ROESY correlation of β-anomeric compounds.

2.3.4. *N*-Glycosylation with Trichloacetimidate Donors

In the course of optimization of the glycoside donor, we investigated trichloroacetimidate glycosides as complements to thioglycosides. In order to test this class of glycoside in *N*-glycosylation of nucleobase, we focused our attention on the identification of a suitable trichloroacetimidate donor. Regioselective hydrolysis of aryl thioglycoside **11** to produce the corresponding lactol was found to be particularly difficult. Indeed, when using an excess of *N*-bromosuccinimide (NBS) in wet EtOAc,[180] only the corresponding lactone **45** could be isolated in 60% yield (**Figure 2.13**).

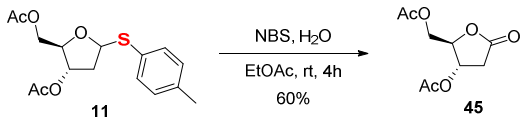


Figure 2.13. Hydrolysis of thioglycoside **11** led to the formation of lactone **45**.

The formation of the lactone can be explained by the mechanistic proposal presented in **Figure 2.14**. The thioglycoside **11** reacted with NBS to form the sulfonium intermediate **A**, which underwent elimination to the oxocarbenium ion **B**. It was then trapped with H_2O to produce the lactol **C**. This intermediate could further react with NBS to form the lactol hypobromite species **D**. Subsequent elimination of HBr furnished the oxidized product **45**.[181] In order to avoid the formation of the lactone, other reaction conditions such as *N*-iodosaccharin (NISac)[182] or the combination of *N*-iodosuccinimide (NIS) and trifluoroacetic acid (TFA)[183] were tested but this resulted in the formation of self-condensation products.

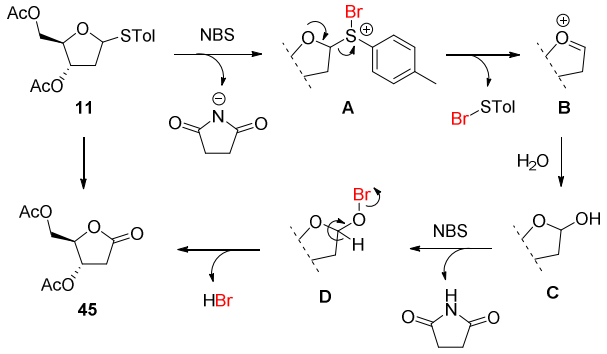


Figure 2.14. Mechanistic proposal for the formation of lactone **45** *via* a one-pot hydrolysis of thioglycoside **11** followed by oxidation of the corresponding lactol.

We then turned our attention to the more stable thioglycoside **14**, which upon hydrolysis with NISac in a mixture of MeCN/H_2O afforded the expected lactol **46** as an anomeric mixture (1.2:1.0) in 72% yield (**Figure 2.15**). Compound **46** was subsequently treated with trichloroacetonitrile (CCl_3CN) and NaH to obtain a clean conversion to the corresponding trichloroacetimidate donor **47**, as observed by ^{1}H-NMR analysis of the crude mixture. This compound was fully characterized (see the Experimental Chapter), and then used in the next step without delay.

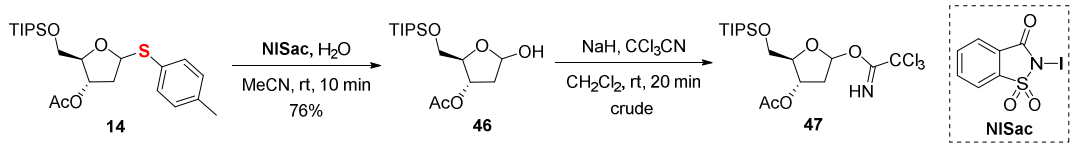


Figure 2.15. Synthesis of trichloacetimidate glycoside **47**.

Imitate **47** was then activated in the presence of TMSOTf and silylated 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21** to undergo *N*-glycosylation (**Figure 2.16**). The corresponding nucleosides **25** were obtained in 69% yield (over two steps from lactol **46**) in a ratio of α/β = 1.0:1.8. This example demonstrated the potential of trichloroacetimidate "Schmidt" glycosyl donors for the synthesis of 2'-β-D-deoxyribonucleosides.

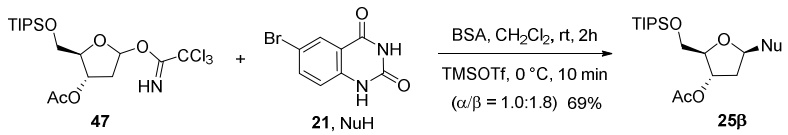


Figure 2.16. *N*-Glycosylation of trichloacetimidate donor **47** with 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21**.

2.4. Mechanistic Considerations for *N*-Glycosylation

2.4.1. Nucleophilic Addition to Five-Membered Ring Oxocarbenium Ions

During 2-deoxyribosylation reactions involving thioglycoside donors, an oxocarbenium intermediate is formed by elimination of the activated thioether aglycon. Evidence for an SN_1 pathway is provided by the enantiopure thioglycoside **14β** (Entry 5, **Table 2.1**), which gives the exact same anomeric ratio of nucleoside products as the 1.0:1.8 (α/β) mixture of thioglycoside **14** (Entry 4, **Table 2.1**). The presence of an oxocarbenium intermediate is further supported by the scope of the glycosylation reaction described in **Table 2.2**, where thioglycoside **14** afforded roughly the same α/β mixtures - independent of the nucleobase structure. Reactions utilizing the common α-glycosyl chloride donor **22α**, in contrast, proceed via SN_2-type pathways, where competing anomerization reactions to the β-glycosyl chloride can result in α-nucleosides as the major product.[154]

In SN_1 reactions involving a 2-deoxyribofuranoside oxocarbenium intermediate, two conformations are possible where C3 is situated either above or below the C2-C1-O-C4 plane, to give the ^3E or E_3

conformer, respectively (**Figure 2.17**).[184-185] The group at C3 principally governs the lowest energy conformer, adopting a pseudoaxial orientation in the case of an alkoxy group. A useful model to explain diastereoselective reactions on such five-membered rings was proposed by Woerpel *et al.*[184-185] where a nucleophile preferentially attacks from the inside face of the oxocarbenium envelope to give the α-product (**Figure 2.17A**). This model explains the stereoselectivity observed in the case of small carbon nucleophiles but it is too simplified to interpret the various outcomes in the case of oxygen or nitrogen nucleophiles.

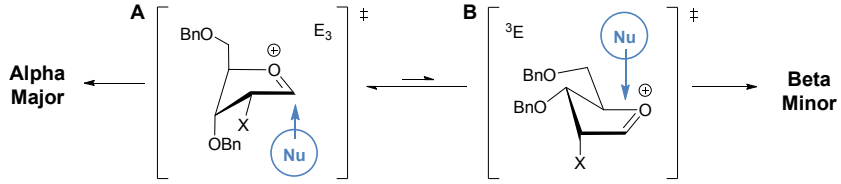


Figure 2.17. Woerpel's inside attack model on five-membered ring oxocarbenium ions. X = OBn, H, F.

In contrast to this model, the NIS/TMSOTf promoted *N*-glycosylation of thioglycoside donors provides primarily the β-anomer. This can be observed with thioglycosides **14**-**16** but in particular with thioglycoside **17**, whose enhanced stereoselectivity is supported by a C1-3-anchimeric participation mechanism (**Figure 2.18**, Intermediate **A**).[175] It is possible that, under our reaction conditions, a counterion occupies the inside face of the oxocarbenium envelope and/or the large steric bulk of the silylated nucleobase causes a change in preference from "inside" to "outside" attack of the nucleophile.

Woerpel's and Ito's groups have reported that conformational perturbations present in fused rings can also mediate stereoelectronic effects.[176,186] For example, when C3-C4 are tethered, only a diequatorial oxocarbenium ion should be accessible (**Figure 2.18**, Intermediate **B**).[186] The resulting ^{3}E conformer gives the opposite stereoselectivity of *N*-glycosylation as compared to the unconstrained thioglycosides, where **18** and **19** give α-anomers as primary products. These results suggest that the selectivity of the reaction is governed by the conformation of the oxocarbenium as well as anchimeric effects.

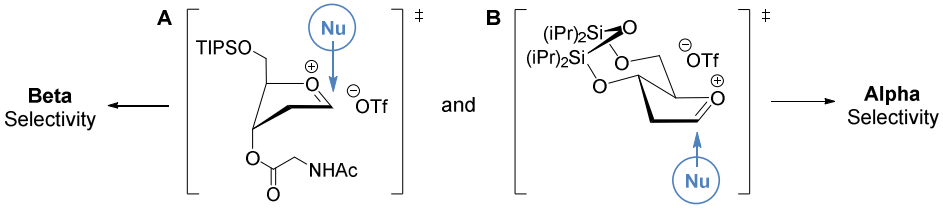


Figure 2.18. Plausible oxocarbenium ions for the preferred formation of α- and β-nucleosides from thioglycosides.

2.4.2. Nucleophilic Addition to Bicyclic Oxocarbenium Ions

Woerpel and co-workers have previously demonstrated that rings fused to five-membered ring oxocarbenium ions exert little influence on the selectivity of nucleophilic addition as long as they are large enough.[186] However, dramatic loss of selectivity could be observed with fused six-membered rings as compared to eight-membered rings (**Figure 2.19**, Intermediates **A** and **B**).[186] In addition, Woerpel *et al.* demonstrated that replacing one carbon by an oxygen atom in the fused seven-membered rings can also largely reduce the selectivity by restricting the conformational freedom of the oxocarbenium.[187] Cycloheptane and oxepane fused to oxocarbenium ions were used to describe the effect on subtle difference in conformational rigidity. Calculations were performed on these two systems separately to evaluate the transition state energy differences resulting from inside and outside attack of nucleophiles (**Figure 2.19**, Intermediates **C**). Very little difference in energy from inside to outside attack was observed for the oxepane-fused system (0.2 kcal mol^{-1}) as compared to the cycloheptane system (1.6 kcal mol^{-1}). This demonstrated the impact of one C-O bond (1.43 Å) as compared to C-C bonds (1.54 Å) on the conformation of fused-ring oxocarbenium ions. This conclusion supports our results with the silylated six- and eight-membered fused-rings thioglycoside that showed reduced and even reversed preference from the inside attack model.

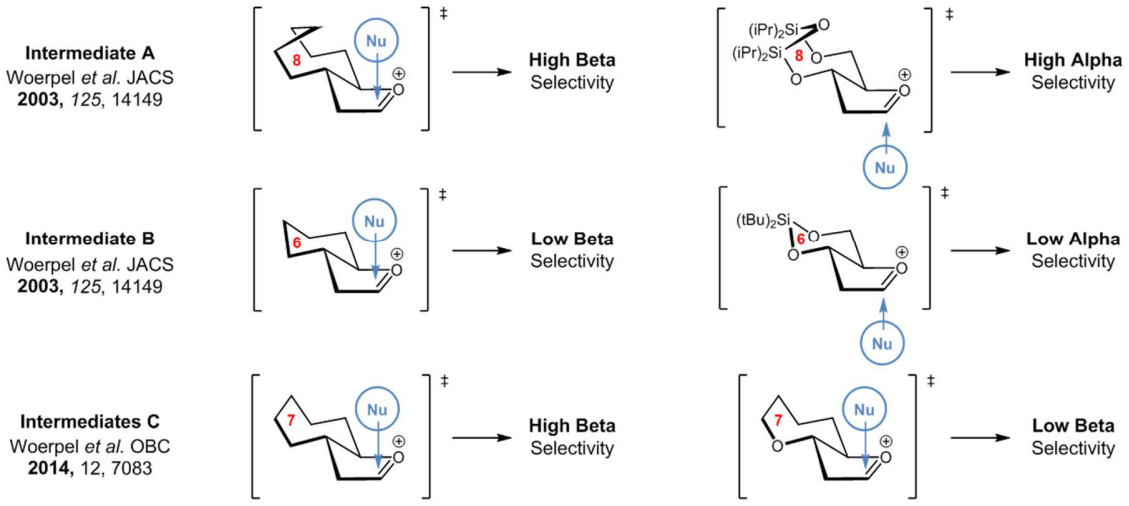


Figure 2.19. Summary of preferred nucleophilic addition to bicyclic oxocarbenium ions. Previously reported data were performed with allyltrimethylsilane as the nucleophile.[186-187]

2.5. Conclusions

Here we report an efficient method for *N*-2-deoxyribosylation of challenging heterocyclic nucleobases. Thioglycosides provide a highly attractive alternative to the commonly used α-glycoside chloride for stereoselective synthesis of β-nucleosides. Thioglycoside donors can be activated in the presence of an *in situ* silylated nucleobase using NIS/HOTf as promoters. By tuning the protecting groups at the C3 and C5 hydroxyls, α/β ratios ranging from 1.0:4.0 to 4.5:1.0 were obtained. This method is compatible with highly challenging expanded nucleobases which were converted into 2'-deoxynucleosides in good yields and β-selectivity. 2-Deoxy-thioriboside coupling reactions tolerate a wide variety of functional groups in the nucleobase structure, including thiophene and thioether groups. Interestingly, all the coupling products where obtained after only 2 minute reaction times with nearly the same α/β selectivity, suggesting a common, highly reactive oxocarbenium intermediate. The stereoselectivity of the reaction was highly influenced by the structure of the thioglycoside. This showed that the nucleophilic attack is driven by the conformation of the oxocarbenium ion in solution, in particular with fused bicyclic system where the conformational rigidity of the oxocarbenium has a dramatic influence on the inside attack model.

As compared to other, more commonly used methodologies, this approach with thioglycoside and imidate donors can provide enhanced yields, β-selectivity, shorter reaction times, and a broader scope of nucleobase substrates. Hence, this provide a powerful means for the synthesis of nucleoside analogues as new fluorescent probes and drug candidates that can expand our current understanding of DNA biology.

Chapter 3 | Solvatochromic Fluorescence of Biaryl Pyrimidine Nucleosides

Chapter 3

Solvatochromic Fluorescence of Biaryl Pyrimidine Nucleosides

3.1. Summary

Fluorescent pyrimidine analogs containing a fused biphenyl unit were prepared in high yields using the previously described stereoselective *N*-glycosylation of 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione followed by Suzuki-Miyaura cross-coupling reactions with various boronic esters. The resulting "push-pull" fluorophores exhibited highly solvatochromic emissions with an exceptionally large positive bathochromic slope (3.7×10^2 cm^{-1} / kcal mol^{-1}) for the dimethylaniline derivative (**Me_2NPh**). This compound is therefore recognized as a promising microenvironment sensitive fluorescent probe. Taken together, the results suggested the presence of twisted intramolecular charge-transfer (TICT) states with the pyrimidine nucleobase acting as an electron acceptor in the excited state.

3.2. Introduction

Fluorescence spectroscopy is one of the most powerful tools in the field of chemical biology.[188-189] Fluorescence-based labeling of nucleic acids is widely used for genomic sequencing,[190] for the detection of infections or genetic diseases,[191] and for the visualization of DNA replication *in vivo*.[100] The non-emissive nature of the natural nucleobases allows the development of fluorescent nucleoside analogs. [109] With their small size and predictable location, fluorescent nucleosides are ideally suited for monitoring biochemical transformations,[130,192] conformational changes,[74] and base pairing interactions.[114,129,155] The environmental sensitivity of these probes allows for reporting of microenvironment parameters such as polarity,[193] viscosity,[194] and pH.[195-196] *C*-Nucleosides containing arene groups such as pyrene and biphenyl can exhibit high quantum efficiencies of emission (**Figure 3.1A**). When incorporated into DNA, they can be used as scaffolds for the construction of multi-chromophore arrays.[197-198] The incorporation of biphenyl *C*-nucleosides into duplex DNA has been shown to result in a zipper-like interstrand stacking recognition motif.[199-200] However, the Watson-Crick hydrogen bonding face in such derivatives is absent, making them poor nucleobase mimics.

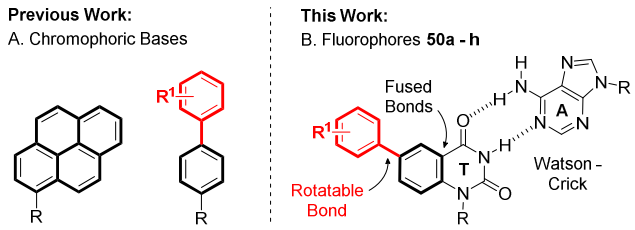


Figure 3.1. (A) Chromophoric C-nucleosides containing pyrene and biphenyl. (B). Emissive biaryl nucleoside analogs **50a** – **h**. R = 2'-deoxyribose, R^1 = electron-withdrawing or donating group.

The molecular structure and fluorescence properties of biphenyls have been widely investigated and it has been shown that the torsion angle is estimated to be 40˚ in the ground state S_0 and nearly planar in the excited state S_1.[201] In particular, donor-acceptor biphenyls exhibited photoinduced charge-transfer (CT) interactions, which was stabilized with increase solvent polarity. This induced the two phenyls to twist and resulted in red-shifted emission and fluorescence quenching.[202-204] These features illustrate the concept of twisted intramolecular charge transfer (TICT) state.[205-206] Given the remarkable photophysical properties of substituted biphenyls,[202-204] and the need for emissive nucleobase mimics,[109] we designed nucleoside analogs **50a** – **h** (**Figure 3.1B**). These new analogs contained a biphenyl unit fused to thymidine in a way that should be compatible with proper Watson-Crick base pairing.[155,207] In such a way, the fluorophores possessed a fused system and a rotatable bond. Herein, we describe the synthesis and evaluation of these new fluorescent nucleobase analogs, and provide a detailed analysis of their structure – photophysical properties relationships.

3.3. Results and Discussion

3.3.1. Synthesis of Biaryl Pyrimidine Nucleosides

The synthesis of the biaryl nucleosides **50a** – **h** was envisioned from the silylated 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione nucleoside **48** *via* a Pd-catalyzed Suzuki-Miyaura cross-coupling (**Figure 3.2** and **Figure 3.3**). In this approach, **48** provides a versatile intermediate for the coupling of various boronic acids and esters. Nucleoside **48** can be obtained from a stereoselective *N*-glycosylation of 2-deoxythioriboside **14** with 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21**. According to our new method described in the previous chapter,[208] 2-deoxythioriboside **14** was *N*-glycosylated with **21** using the

combination of *N*,*O*-bis-(trimethylsilyl)acetamide (BSA), *N*-iodosuccinimide (NIS) and trimethylsilyl trifluoromethane-sulfonate (TMSOTf) to afford **25** in 84% yield (**25α**/**25β** = 1.0:1.8, 1.0 g scale, **Figure 3.2**). The reaction was also performed on a larger scale (>10 g scale, single batch) to obtain **25β** without significant loss of isolated yield (82%).

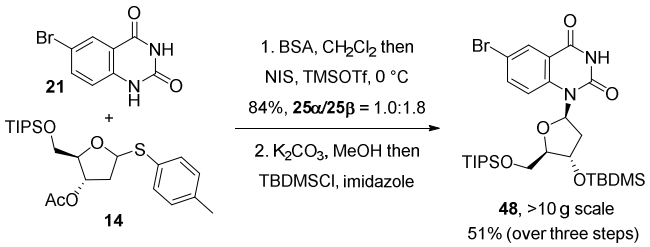


Figure 3.2. Stereoselective *N*-glycosylation of 2-deoxythioriboside **14** and synthesis of intermediate **48**.

To facilitate cross-coupling reactions, the C'3-acetate of **25β** was converted into a C'3-silyl ether to give **48** in 95% yield over two steps (**Figure 3.2**). The influence of the protecting groups on the sugar moiety in Pd-catalyzed cross-coupling has already been examined. It has been shown that the bulky TBDMS group prevents coordination of the Palladium catalyst.[209] Suitable conditions for Pd-catalyzed Suzuki-Miyaura cross-coupling between **48** and phenyl boronic acid were evaluated by screening various palladium sources, ligands, bases and solvents. Test reactions indicated that $Pd(dppf)Cl_2$ is a superior catalyst as compared to $Pd(PPh_3)_2Cl_2$, $Pd_2(dba)_3$/SPHOS and $Pd(OAc)_2$/SPHOS. Furthermore, dioxane was found to be superior to toluene and DMF. The combination of $Pd(dppf)Cl_2$ and AcOK in dioxane at 90 °C were found to be highly effective for this transformation. The biphenyl product **49a** ("**Ph**") was isolated in 91% yield (**Figure 3.3**). Electron-donating and withdrawing groups on the phenyl boronic acid were well tolerated, giving good to excellent yields for biphenyl derivatives **49b** – **d** ("**MeOPh**", 90%), ("**Me$_2$NPh**", 67%, performed on a 4.0 g scale), and ("**CNPh**", 91%). This approach is also compatible with the synthesis of derivatives containing five-membered heteroaromatic groups **49e** – **h**. It is known that five-membered, 2-heteroaromatic boronic acids are prone to undergo protodeboronation under basic conditions,[210-211] but reactions involving **48** and thiophene, pyrrole or furan boronic acids furnished the corresponding biaryls **49e** – **h** in good isolated yields ("**2Th**", 62%), ("**2MePy**", 82%), ("**2Fu**", 64%), and ("**3Fu**", 70%).

Treating nucleosides **49a** – **h** with TBAF in THF afforded the deprotected compounds **50a** – **h** in isolated yields ranging from 70% to 92% (**Figure 3.3**), and in high purity of >98% according to ^{1}H-NMR.

In addition to column chromatography, extensive washing of the products with MeOH and H_2O/AcOH was used to eliminate traces of TBAF salts. Copies of ^1H and ^{13}C NMR spectra are available in the Supporting Information of the published work.[212]

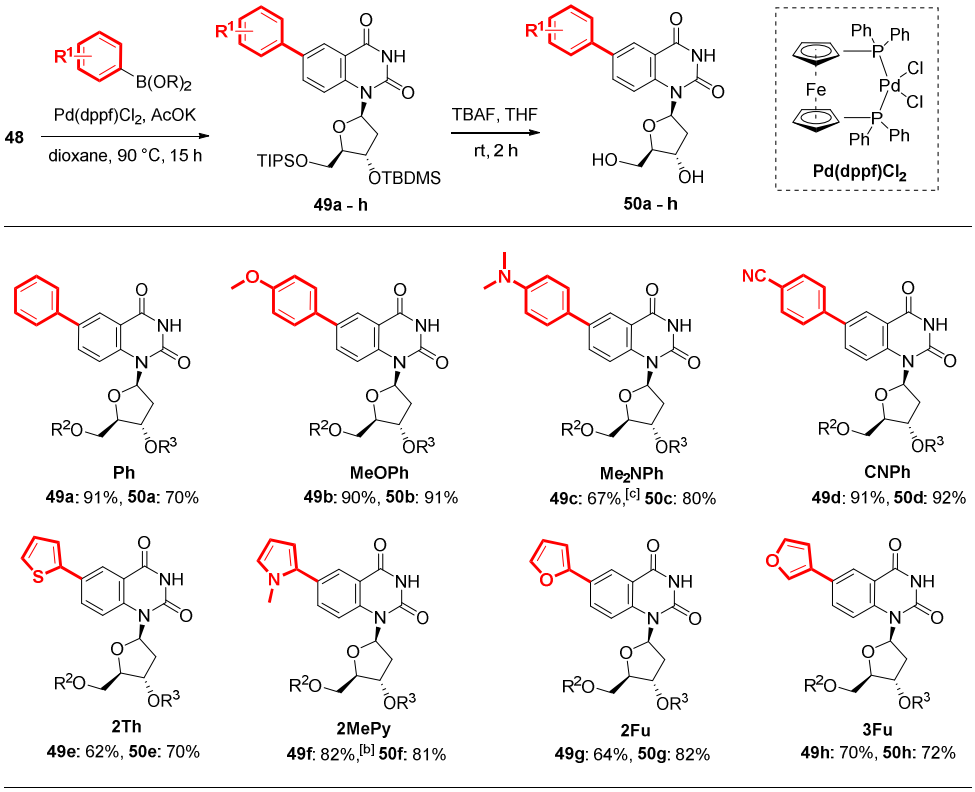


Figure 3.3. Pd-Catalyzed Suzuki-Miyaura cross-coupling and silyl groups deprotection. Reagent and conditions: [a] **48** (1.0 equiv), Ar-B(OR)$_2$ (1.5 equiv), AcOK (1.5 equiv), Pd(dppf)Cl$_2$ (0.05 equiv), dioxane (0.07 M), 90 °C, 15 h. [b] Ar-B(OR)$_2$ (2.0 equiv), and AcOK (2.0 equiv) were used. [c] The reaction was performed on a 4.0 g scale. **49a – h**: R^2 = TIPS, R^3 = TBDMS; **50a – h**: $R^2 = R^3$ = H.

3.3.2. Photophysical Properties of the Nucleosides

The photophysical properties of **50a – h** are summarized in **Table 3.1**. All reported derivatives exhibited large extinction coefficient values in water ($\varepsilon_{276\text{-}300} = 1.87 - 3.03 \times 10^4$ M^{-1} cm^{-1}) as compared to thymidine-5'-monophosphate ($\varepsilon_{260} = 0.86 \times 10^4$ M^{-1} cm^{-1}).[213] The quantum yields of these compounds are moderate in water ($\phi = 0.14 - 0.001$) and in dioxane ($\phi = 0.21 - 0.07$). These values are very similar to biphenyl itself ($\phi = 0.18$),[214] but the wavelengths of excitation and emission of **50a – h** are red-shifted. For example, the maximal wavelengths of absorbance and emission of the biphenyl nucleoside **Ph** (λ_{abs} = 270 nm and λ_{em} = 371 nm) are 23 nm and 59 nm to the red of biphenyl in dioxane (λ_{abs} = 247 nm and λ_{em}

= 312 nm). The reason for this red-shifted emission is related to the ability of the pyrimidine group to serve as an electron accepting group upon photoexcitation. DFT calculations indicate charge transfer from the benzene group of **Ph** to pyrimidine upon going from HOMO → LUMO (**Figure 3.4**).

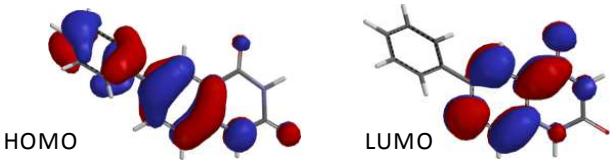


Figure 3.4. Molecular orbital plot of the isolated nucleobase of **50a**. Calculations were performed at the B3LYP level using Spartan 10.

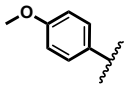
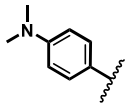
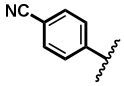
Substituent	Name	Solvent	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ε (λ_{max})[d]	ϕ[e]
	**Ph**	dioxane	270	371	10.1		0.11
	(**50a**)	water	269	404	12.4	2.49	0.085
	**MeOPh**	dioxane	280	390	10.1		0.132
	(**50b**)	water	276	462	14.6	2.74	0.035
	**Me_2NPh**	dioxane	312	459	10.3		0.071
	(**50c**)	water	297	460	11.9	2.01	0.010
	**CNPh**	dioxane	293	362	6.50		0.206
	(**50d**)	water	286	388	7.40	2.07	0.140
	**2Th**	dioxane	298	391	8.00		0.094
	(**50e**)	water	295	449	11.6	2.02	0.024
	**2MePy**	dioxane	292	432	11.1		0.079
	(**50f**)	water	286	375	8.30	1.92	0.001
	**2Fu**	dioxane	292	403	9.40		0.152
	(**50g**)	water	290	463	12.9	3.03	0.038
	**3Fu**	dioxane	269	381	10.9		0.086
	(**50h**)	water	267	430	14.2	1.87	0.075

Table 3.1. Photophysical data of emissive nucleosides **50a – h**. [a] λ_{abs} are reported at the maximum absorbance wavelength in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1}. [d] Extinction coefficients (ε) are given at λ_{max} in H_2O and are reported in 10^4 M^{-1} cm^{-1}. [e] 2-Aminopyridine in 0.1 N H_2SO_4 (ϕ = 0.60) was used as the fluorescent standard for the relative quantum yields (ϕ) of **50a – h**.[215] (See **Appendix A1 – A6** for complete fluorescence data, spectra and plots).

Taken together, DFT calculations and photophysical properties of compounds **50a – d** suggest the presence of a push-pull system, where the pyrimidine group acts as an electron acceptor in the excited state. The magnitude of this push-pull effect is augmented by electron donating groups in **50b** and **50c**, but diminished by the electron withdrawing cyano group in **50d**. Indeed, the calculated HOMO-LUMO gaps (**Me_2NPh** < **MeOPh** < **Ph** < **CNPh**) were inversely correlated with the Stokes shifts of these compounds in dioxane (**Table 3.1**). Taken together, these results are consistent with the presence of emissive TICT states that are stabilized by electron-donating groups on the benzene ring.

3.3.3. Twisted Intramolecular Charge Transfer (TICT) State

Charge separation processes in organic fluorophores occur favorably in the case of twisted conformation of two moieties, composed of an electron rich (donor) and an electron poor (acceptor) system (**Figure 3.5**). This process results in a twisted intramolecular charge transfer (TICT) emissive state.[216]

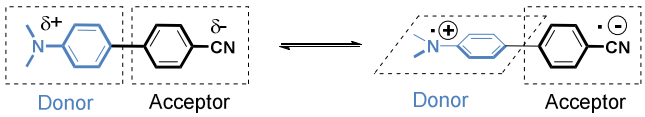


Figure 3.5. Structural representation of the change of geometry upon going from the normal emissive state to the charge transfer (TICT) state for a simple fluorescent donor – acceptor biphenyl.[202]

The TICT state can be characterized by two main properties, which is the charge separation and the twisted geometry. The charge transfer amount for one electronic charge (in the order of 0.8) from the donor to the acceptor.[217] The resulting radical ion pair is highly polar and is therefore stabilized by a polar environment. Experimentally, fluorescent molecules having TICT states exhibit red-shifted emission bands as the polarity of solvent increases.[202] The twisted geometry brings the donor and the acceptor almost perpendicular to one another and are therefore orbitally decoupled. This results in a forbidden overlap of the perpendicular orbitals of the two π-systems. TICT states are therefore normally characterized with low fluorescent quantum yields.[218] The structure of the designed nucleosides **50a – h** are able to adopt a twisted conformation and the photophysical properties of the **Me_2NPh** derivative are consistent with the presence of a TICT emissive state (**Figure 3.6**).

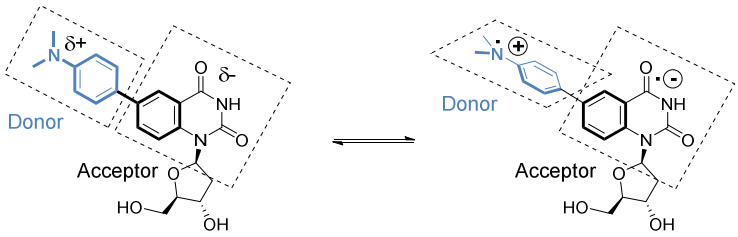


Figure 3.6. Structural representation of the change of geometry upon going from the normal emissive state to the TICT emissive state of compound **50c**.

3.3.4. Solvatochromic Fluorescence

To characterize the environmental sensitivity of nucleosides **50a** – **h**, the maximal wavelengths for absorbance and emission were measured in different organic solvents or water/dioxane mixtures, and the Stokes shifts were plotted against Reichardt's solvent polarity parameter (E_T^{30}).[219-220] **Table 3.2** describes the photophysical properties of **Me$_2$NPh** upon increasing solvent polarity.

Solvent	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]	E_T^{30}[e]
dioxane	312	459	10.3	0.071	36.63
THF	312	467	10.6	0.053	37.40
EtOAc	312	482	11.3	0.048	38.10
$CHCl_3$	318	503	11.6	0.049	39.10
CH_2Cl_2	317	513	12.1	0.049	40.70
DCE	316	516	12.3	0.051	41.30
DMF	315	536	13.2	0.020	43.20
DMSO	316	549	13.4	0.013	45.10
MeCN	312	550	13.9	0.008	45.60

Table 3.2. Photophysical data of **Me$_2$NPh** in organic solvents. [a] λ_{abs} are reported at the maximum absorbance wavelength in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1} [d] 2-Aminopyridine in 0.1 N H_2SO_4 (ϕ = 0.60) was used as the fluorescent standard for the relative quantum yields (ϕ). [e] Reichardt's solvent polarity values were obtained from the literature.[219]

Having a constant twisted structure in the ground state, the absorption maxima of **50a** – **h** were found to be relatively similar in the different solvents (λ_{abs} = 312 – 317 nm). Previous studies have shown that the first excited state of these donor-acceptor biphenyls is independent of the twist angle and the solvent

polarity, whereas the emissive state is highly sensitive to these parameters.[202-204] Consistent with the presence of a twisted biaryl bond in the ground state of these molecules, the absorption maxima of **50a** – **50h** were relatively insensitive to solvent polarity. The Stokes shifts of these compounds are highly solvatochromic due to changes in their emission wavelengths (**Figure 3.7A**). For example, the Stokes shift of **Me$_2$NPh** increased from 10.3 × 10^3 cm^{-1} in dioxane (λ_{em} = 459 nm) to 13.9 × 10^3 cm^{-1} in MeCN (λ_{em} = 550 nm) to give an exceptionally large slope of 373 cm^{-1} / kcal mol^{-1} (**Figure 3.7B**). This effect can easily be observed under UV illumination, where samples of **Me$_2$NPh** appear bright blue to dark orange, depending on the solvent polarity.[212] The emission intensity of **Me$_2$NPh** is highly sensitive to solvent polarity, with a slope of 0.006 mol kcal^{-1} from ϕ = 0.071 in dioxane, to ϕ = 0.008 in MeCN (**Figure 3.7C**). **Me$_2$NPh** exhibits low quantum yield in water (ϕ = 0.010) that is 2.7-fold higher in D_2O (ϕ = 0.027). This type of kinetic isotope effect suggests the presence of proton-transfer reactions between the excited fluorophore and bulk solvent that can facilitate non-emissive decay in water.[221] Indeed, in polar media, the solvation stabilizes the transfer of electron from the donor to the acceptor and subsequently twist the donor. The resulting localization of the two orbitals perpendicular to each other on different π-systems involve during transition from the TICT state to the ground state results in a low fluorescence quantum yield.[216]

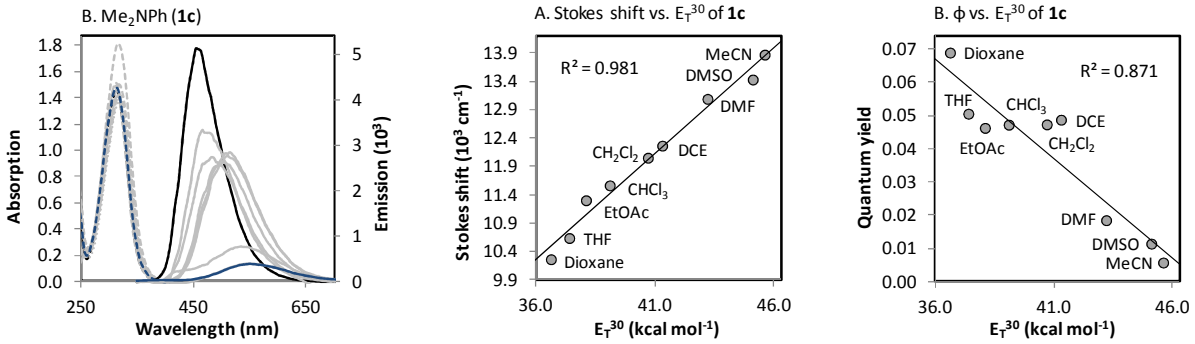


Figure 3.7. A) Absorbance (---) and emission (—) spectra of **Me$_2$NPh** in water (blue), dioxane (black) and mixtures (grey). B) Plot between the Stokes shift of **50c** and E_T^{30} value. C) Plot between the quantum yield of **1c** and E_T^{30} value.

Consistent with the presence of emissive TICT states, the emissions from **50a** –**h** exhibit red-shifting and lower quantum yields with increasing solvent polarity. The magnitudes of these effects are compared in **Figure 3.8**. The biphenyl derivatives containing electron-donating substituents **MeOPh** and **Me$_2$NPh** exhibit highly solvatochromic Stokes shifts and quantum yields. In contrast, **CNPh** exhibits diminished

solvent sensitivity, and a good quantum yield over the entire range of solvent polarity tested (ϕ = 0.21 – 0.14). Together these results are consistent with "push-pull" TICT fluorophores. The trends between the compounds containing a 5-membered heterocyclic ring **50e** – **h** are somewhat more complicated. For example, **2Fu** exhibits quantum yields that are highly sensitive to solvent polarity (4.2×10^{-3} mol kcal^{-1}), whereas it isomer, **3Fu**, exhibits very little sensitivity (0.3×10^{-3} mol kcal^{-1}).

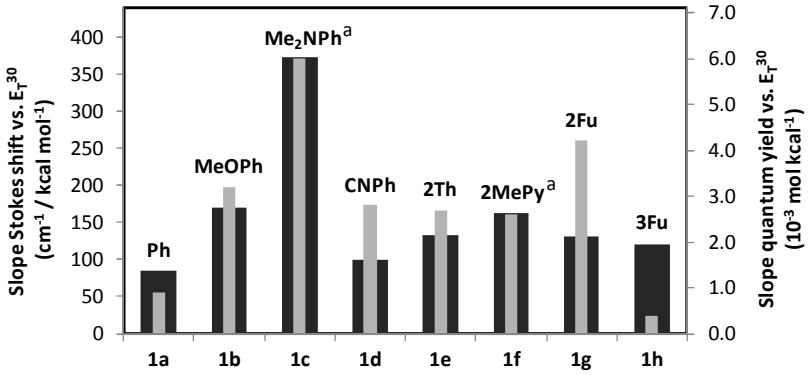


Figure 3.8. Sensitivity of the nucleosides **50a – h** towards solvent polarity in dioxane, water and their mixtures. The black bars represent the slope of the Stokes shift and E_T^{30} value. The grey bars represent the slope of the fluorescence quantum yield and E_T^{30} value. [a] Due to quenching by H_2O, slopes were determined in organic solvents for **50c** and **50f**.

To systematically compare the theoretical and experimental photophysical properties of the biphenyl derivatives **50a** – **d** with the biaryl derivatives **50e** – **h**, the calculated HOMO – LUMO gap for each compound was plotted against its peak emission wavelength in dioxane to give an excellent linear correlation (R^2 = 0.981, **Figure 3.9A**). These results suggested that the underlying photophysical processes are similar for **50a** – **h**, where electron-donating groups cause smaller HOMO – LUMO gaps and red-shifted emissions as compared to electron withdrawing groups.

3.3.5. Nucleophilicity Parameters and Fluorescence

The nucleophilicity and electrophilicity of chemical entities can be scaled in order to predict chemical reactions.[222] The most comprehensive scales characterize electrophiles by one solvent-independent parameter (E), and nucleophiles by two solvent-dependent parameters, the nucleophilicity parameter (N) and the sensitivity parameter s_N.[223] Plotting the emission wavelength of each nucleoside versus the Mayr's nucleophilicity parameter (N) for the isolated aryl/heteroaryl group (R^1),[224-228] gave

an excellent linear correlation between these factors (R^2 = 0.983, **Figure 3.9B**). This demonstrated that similar photophysical processes are responsible for the fluorescence of all nucleoside analogs **50a** – **h**. This unusual approach reveals a possible method to extrapolate nucleophilicity parameters of groups that can not be experimentally measured. For example, this plot can be used to predict a nucleophilicity parameter for the position 3 of furan ($N \approx$ - 4.12) based upon the emission wavelength of **3Fu** (λ_{em} = 388 nm) in CH_2Cl_2.

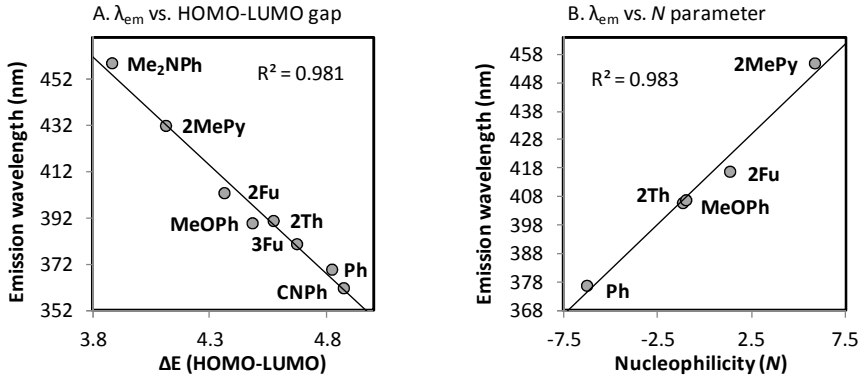


Figure 3.9. (A) Plot between the emission wavelength of **50a** – **h** and calculated HOMO-LUMO energy gaps. See Experimental Section for calculation details. (B) Plot between the emission wavelength of the nucleosides **50a** – **h** and Mayr's nucleophilicity parameter (N) of the isolated aryl groups (R^1) for benzene, anisole, thiophene, *N*-methylpyrrole, and furan in CH_2Cl_2.[224-228]

3.4. Conclusions

In summary, we have synthesized a new family of fluorescent biaryl pyrimidines in high yield and purity. The nucleosides exhibited highly solvatochromic emissions from twisted intramolecular charge-transfer (TICT) states. Previous studies have demonstrated fluorescent emissions resulting from charge-transfer recombination in pyrene-modified deoxyuridine,[229] but to the best of our knowledge, derivatives **50a** – **h** provided the first examples of nucleobase analogs that are classic "TICT" fluorophores. In addition, **Me_2NPh** (**50c**) exhibited absorbance well to the red of unmodified DNA nucleobases and emissions that are exceptionally sensitive to solvent polarity. **Me_2NPh** is therefore a promising candidate for investigating local DNA conformational changes by reporting changes in microenvironmental polarity and solvation. In addition of providing new fluorescent nucleoside analogs, the study described here gives a basis for the design and implementation of other fluorescent probes, based on the remarkable properties of "push-pull" systems.

Chapter 4 | A Fluorescent Cytosine for Proton-Coupled DNA Folding

Chapter 4
A Fluorescent Cytosine for Proton-Coupled DNA Folding

4.1. Summary

Ionized nucleobases participate in pairing interactions outside of Watson and Crick's rules. Base pairing and ionization can be coupled via global conformational changes to raise the apparent pK_a of protonated nucleobases to values above physiological pH. To provide the first specific reporter of proton-coupled DNA folding, we developed a "push-pull" fluorescent nucleoside analog composed of dimethylaniline (DMA) fused to deoxycytidine. "^{DMA}**C**" exhibits the same pK_a and base paring characteristics as native cytosine residues in the human telomeric repeat sequence, where it causes little or no perturbation of DNA structure or stability. Upon protonation of ^{DMA}**C**, enhanced charge transfer results in large redshifting (+40 nm) of its excitation/emission maxima (**Figure 4.1**).

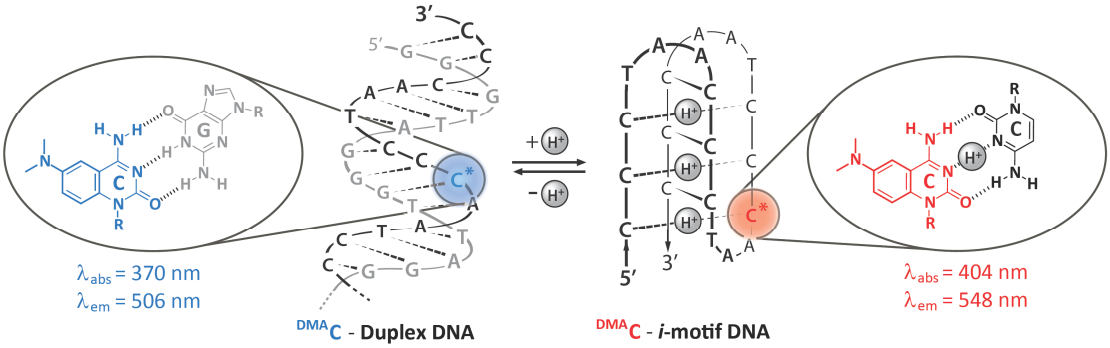


Figure 4.1. A fluorescent cytosine mimic (^{DMA}**C**) for proton-coupled DNA folding.

^{DMA}**C**'s fluorescence intensity, anisotropy, and energy transfer properties can be used to track conformational changes in real time. Strand-displacement assays were conducted by mixing ^{DMA}**C**-labeled duplexes containing a 5' single-stranded overhang with an excess of unlabeled DNA to initiate thermodynamically favorable unfolding-refolding reactions that release the ^{DMA}**C**-labeled strand from its complement (**Figure 4.2**). Rate constants for strand displacement upon addition of *i*-motif DNA (k = 1.0 M^{-1} s^{-1}, $t_{½}$ ≈ 12 hours) were 320-fold lower than those measured upon addition of unfolded DNA (k = 3.2 × 10^2 M^{-1} s^{-1}, $t_{½}$ ≈ 2 minutes). These results reveal that *i*-motif structures having marginal thermodynamic stabilities (T_m < 40 °C) can still pose large kinetic barriers to duplex formation under near-physiological conditions of pH (5.75), temperature (25 °C), and salt (100 mM NaCl).

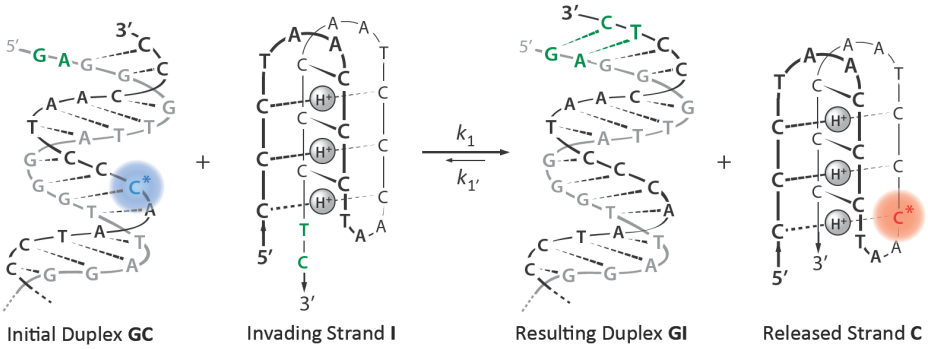


Figure 4.2. Strand-displacement assays revealed slow exchange between *i*-motif and duplex DNA.

4.2. Introduction

Acid-loving bacteria can thrive in strongly acidic conditions (pH $\leq$ 2) by maintaining their intracellular pH in a range of 4.6 – 6.5.[230] The nuclear pH of eukaryotic cells is typically 7.3,[39] where the canonical nucleobases are essentially uncharged.[31-32] Ionized nucleobases are nonetheless observed in certain DNA and RNA structures at neutral pH.[231-238] Long-range electrostatic and local hydrogen bonding interactions can raise the effective pK_a (pK) of protonated A, C, and G residues to values >7.0 in folded structures.[231-236] Reversible nucleobase protonation can therefore mediate complex biological functions, such as translational framesshifting,[237] and acid-base catalysis.[238]

Cytidine is the most intrinsically basic canonical nucleobase, with a pK_a = 4.5 for its conjugate acid (C^+) in water.[31-32] G-C base pairs in duplex DNA exist in two forms that rapidly interconvert between Watson/Crick and transient G-C^+ Hoogsteen base pairs that exhibit an apparent pK $\approx$ 7.2.[239] G-C^+ "wobble" base pairs have also been observed at pH $\approx$ 7.0 in disrupted DNA and RNA helicies.[230,240] CG-C^+ triplets involving both Watson and Crick and Hoogsteen base pair interactions can exhibit pK values as high as 9.5.[236]

The "intercalated" C-C^+ base pairs of *i*-motif structures are well characterized *in vitro* (**Figure 4.3**),[20,24] but their potential biological relevance remains an open question.[45,48-49,241] Proton-coupled folding of *i*-motif DNA is a highly cooperative process, with complete folding transitions occurring over as little as 0.3 pH units.[242] The apparent pK values for cytosine residues in *i*-motif structures are typically in the range of 5.5 – 6.6,[34] but these values can reach 7.5 or higher as a result of partial DNA dehydration

and/or ligand binding.[35-36,243] While much is known about the thermodynamics of *i*-motif folding *in vitro*,[12] very little is known about the kinetics of folding pathways involving competition between *i*-motif and duplex structures.[244] This competition might influence chromosome stability and/or gene expression. Recent studies by Hurley and co-workers have demonstrated that ligand-mediated stabilization of *i*-motif structures can impact the expression of *BCL2* and its associated proteins.[50-51] However, the unambiguous detection of *i*-motif folded structures, or any other nucleic acid structure containing C^+ in cells remains elusive.

DNA sequences present in large copy numbers provide attractive opportunities for the detection of non-canonical base pairing interactions in cells.[15,245] We therefore selected the cytosine rich human telomeric repeat sequence $({}^{5'}\text{CCCTAA}^{3'})_n$ for developing the first fluorescent reporter of proton-coupled DNA folding.[22,246-248] Unlike traditional biophysical techniques,[107] fluorescence spectroscopy has the unique advantage that it is widely applicable in cell-based assays.[104-105] FRET analysis of end-labeled oligonucleotides is commonly used to study *i*-motif folding,[106,249] but the attachment of two large fluorophores can perturb both the kinetics and thermodynamics of folding reactions.[124-125] An alternative approach could involve the utilization of "internal" fluorescent probes that mimic the structure and function of endogenous nucleobases.[74-75,109,129,155,192,250-251] A handful of fluorescent cytosine analogs have been previously reported,[115-116,127,170,252-253] and some have been incorporated into the loops of *i*-motif structures,[157,254-255] but no probe has been previously incorporated directly into C-C^+ base pair.

Here we report the design and synthesis of a new fluorescent cytidine mimic that can accurately and innocently report cytosine ionization upon proton-coupled DNA folding. $^{\text{DMA}}\mathbf{C}$ has little or no impact on the structure or stability of duplex or *i*-motif structures containing it, where it exhibits the same ionization characteristics as native cytosine residues. We have used $^{\text{DMA}}\mathbf{C}$-modified DNA in strand-displacement experiments to track the interconversion of duplex and *i*-motif structures under constant conditions of pH, salt and temperature. Interestingly, *i*-motif structures exhibit only marginal thermodynamic stabilities ($T_m \approx 38$ °C) at pH = 5.75, but still pose a large kinetic barriers ($t_{1/2} \approx 12$ hours) to duplex formation at 25 °C. In addition to providing the first fluorescent reporter of proton-coupled folding for nucleic acids, these results suggest that energy-dependent processes would be needed to resolve *i*-motif structures on a reasonable time scale *in vivo*.

4.3. A Fluorescent Cytosine Mimic

4.3.1. Probe Design

The new fluorescent nucleoside analog "*N*,*N*-dimethylaniline-2'-deoxycytidine" (^{DMA}**C**) is composed of a dimethylaniline unit ring-fused to 2'-deoxycytidine (**Figure 4.3**). The resulting quinazoline system was selected for its useful fluorescence properties,[212,256-257] and for its ability to be well accommodated in duplex DNA and RNA.[114,130,207] Using high resolution structures,[26-27] molecular modeling studies suggested that a quinazoline-based cytidine mimic would cause no steric clashes in duplex or *i*-motif structures, even when it is further modified at position *C6* or C7. To facilitate sensitive reporting of *N3* protonation, an electron-donating dimethylamino group was included at the *C6* position. This position was selected because it is not in conjugation with *N3*, and therefore expected to have little impact on its basicity. Despite the lack of conjugation between *C6* and *N3*, DFT-based HOMO-LUMO calculations suggested a shift in electron density from dimethylaniline to cytosine upon photoexcitation of ^{DMA}**C** that is enhanced upon protonation of *N3* (**Figure 4.4**). The calculated HOMO-LUMO gap for *N3*-protonated ^{DMA}**C** (ΔE = 2.86 eV) is lower than that of neutral ^{DMA}**C** (ΔE = 3.75 eV). This suggested a red-shift in fluorescence emission would occur upon ^{DMA}**C** protonation at *N3*. In contrast, protonation at the *N6* dimethylamino group should result in blue shifted emission, due to its larger HOMO-LUMO gap (ΔE = 4.27 eV, **Figure 4.4**).

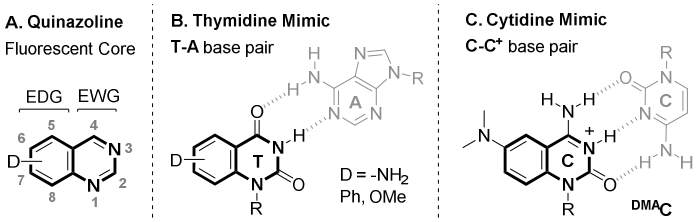


Figure 4.3. (A) Structure and numbering of quinazoline.[212,256-257] (B) Quinazoline-based thymidine mimic in Watson-Crick T-A base pair.[114,130,207] (C) Cytosine mimic ^{DMA}**C** in hemi-protonated C-C$^+$ base pair. R = deoxyribose.

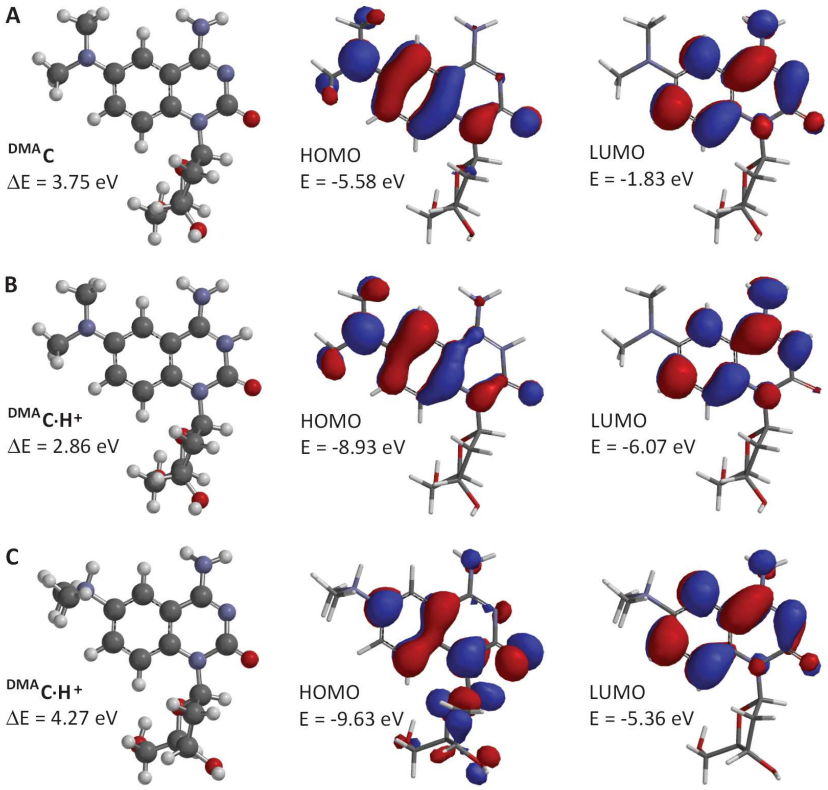


Figure 4.4. HOMO and LUMO molecular orbitals of (A) ^{DMA}**C** (B) ^{DMA}**C**·H$^+$ protonated at *N3* and (C) ^{DMA}**C**·H$^+$ protonated at *N6* obtained from DFT-optimized geometries using B3LYP/6-311++G**.

4.3.2. Synthesis and Solution Conformation of ^{DMA}C

Synthesis commenced from a previously reported 6-bromoquinazoline-2,4-(3*H*)-dione nucleoside **48** (**Figure 4.5**).[208] Pd-catalyzed Buchwald-Hartwig amination of aryl halide **48** by Me_2NH was performed using a combination of $Pd_2(dba)_3$ and JohnPhos[258] to afford nucleoside **51** in 82% yield. To convert the resulting thymidine analog into a cytidine analog, **51** was treated with 2,4,6-triisopropylbenzene-sulfonyl chloride (TPSCl) followed by aqueous ammonia. This two-step procedure afforded nucleoside **52** in 75% yield, which was deprotected with TBAF to give the ^{DMA}**C** nucleoside (**53**) in a 91% isolated yield.

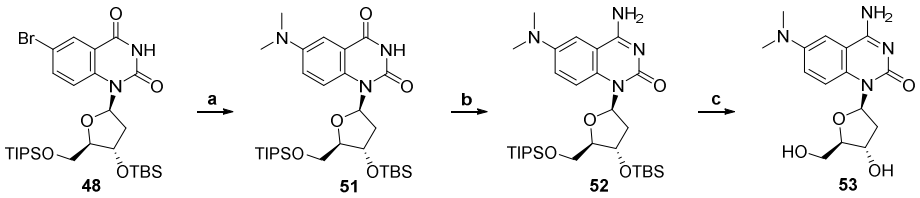


Figure 4.5. Synthesis of ^{DMA}**C** nucleoside (**53**). Reaction conditions: a) Me_2NH in THF (2.0 M), $Pd_2(dba)_3$ (5 mol%), JohnPhos (20 mol%), KOtBu, dioxane, (0.15 M) 60 °C, 2 h, 82% yield. b) (i) TPSCl, Et_3N, DMAP, CH_2Cl_2, 23 °C, 12h, (ii) NH_4OH aq., THF, 23 °C, 10 min, 75% yield (over two steps). c) TBAF, THF, 23 °C, 2 h, 91% yield.

^{1}H-^{1}H ROESY spectra of nucleoside **53** indicate that the glycosidic bond adopts an *anti*-conformation. Strong cross-peaks between H8 and H2'β as well as H8 and H3' were observed (**Figure 4.6A**). These are consistent with atomic distances obtained from a DFT energy-minimized *anti*-conformation (H8–H2'β = 2.39 Å and H8–H3' = 2.90 Å, **Figure 4.6B** and **Table 4.1**). In contrast, the DFT energy-minimized *syn*-conformation exhibits larger atomic distances (H8 – H2'β = 4.77 Å and H8 – H3' = 5.67 Å, **Table 4.1**). Additional ^{1}H-^{1}H ROESY correlations between H1' – H2'α and H1' – H4 indicate a β-configuration at *C1*. The ^{DMA}**C** nucleoside therefore possesses the same stereochemistry and conformational preferences as native deoxycytidine.

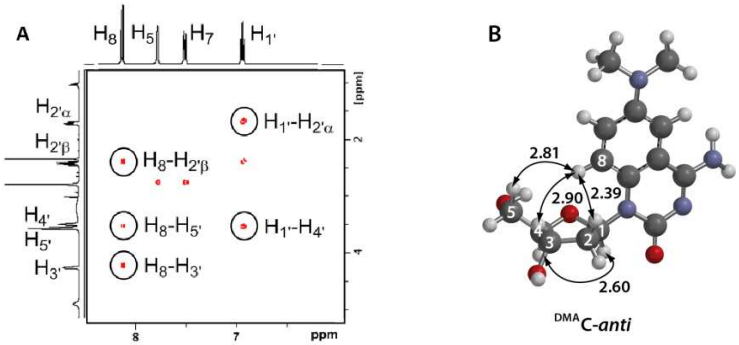


Figure 4.6. Evaluation of glycosidic bond conformation. (A) Partial 2D ROESY of ^{DMA}**C** nucleoside (**53**) in DMSO-d_6 (500 MHz, mixing time of 250 ms). (B) DFT-optimized geometry of the *anti* conformer of ^{DMA}**C** with bond distances in angstroms.

A	Hydrogen	^{DMA}**C**-*anti*	^{DMA}**C**-*syn*
	H8-H2'α	3.844	4.194
	H8-H2'β	2.393	4.771
	H8-H3'	2.905	5.670
	H8-H4'	4.086	4.580
	H8-H5'	2.810	5.910

B	Hydrogen	^{DMA}**C**-*anti*	^{DMA}**C**-*syn*
	H1'-H2'α	2.339	2.354
	H1'-H2'β	3.028	3.026
	H1'-H3'	3.893	3.838
	H1'-H4'	2.603	3.443
	H1'-H5'	4.552	4.395

Table 4.1. Inter-atomic distances (Å) for the *anti* and the *syn* conformers of nucleoside ^{DMA}**C** obtained from DFT-optimized geometries. (A) Hydrogen-hydrogen distances between H8-HX to determine the conformation of the glycosidic bond and (B) hydrogen distances between H1-HX to determine the anomeric configuration at C1.

4.3.3. Photophysical Properties of ^{DMA}C Nucleoside

The maximal absorbance wavelengths of **53** (λ_{abs} = 365 – 425 nm, **Table 4.2**) are more than 100 nm to the red of unmodified nucleobases, allowing for selective excitation in the context of modified DNA. Depending on the solvent polarity and acidity, ^{DMA}**C** exhibits highly variable quantum yields (ϕ = 0.03 – 0.36) and maximal wavelengths of fluorescence emission (λ_{em} = 462 – 526 nm).

Solvent	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ε_{260}[d]	ε_{abs}[d]	ϕ[e]
water	365	526	8.4	17.0	2.5	0.03
dioxane	390	462	4.0	20.0	2.9	0.36
MeCN	395	480	4.5	22.1	2.6	0.23
MeCN [f]	425	555	5.5	27.3	3.4	0.03

Table 4.2. Photophysical properties of ^{DMA}**C** nucleoside (**53**). [a] λ_{abs} (nm) are reported for the most red-shifted absorbance maxima [b] λ_{em} in nm. [c] Stokes shifts in 10^3 cm^{-1}. [d] Extinction coefficients (ε) in 10^3 M^{-1} cm^{-1}. [e] Quinine hemisulfate (ϕ = 0.55 at λ_{ex} = 370 nm) in 0.5 M H_2SO_4 was used as the fluorescent standard for the relative quantum yields (ϕ) of ^{DMA}**C**.[259] [f] 1.0 equiv. of anhydrous HCl was included.

To systematically characterize the environmental sensitivity of ^{DMA}**C**, absorption and emission spectra were measured in water/dioxane mixtures (**Figure 4.7** and **Table 4.3**). By plotting Stokes' shifts against Reichardt's solvent polarity parameter (E_T^{30}, **Table 7.1** in the Experimental Chapter) a linear slope of 166 cm^{-1} / kcal mol^{-1} (R^2 = 0.987) was observed.[219-220] This large slope validates ^{DMA}**C** as a push-pull fluorophore.[229,260]

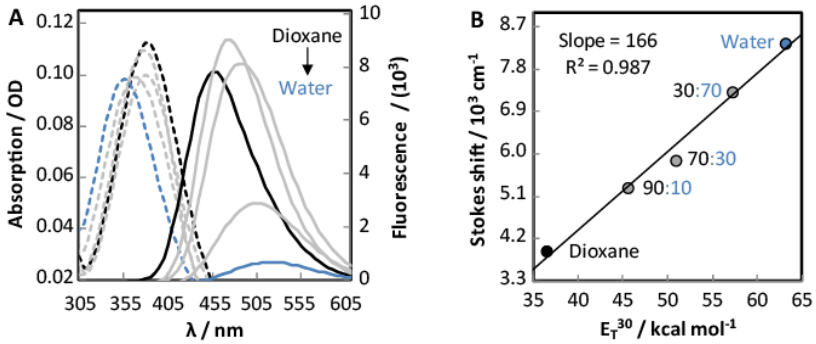
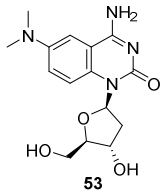


Figure 4.7. (A) Absorbance (---) and emission (—) spectra (λ_{ex} = 370 nm) of ^{DMA}**C** in dioxane (black), water (blue) and mixtures (grey). (B) Linear correlation between the Stokes shift and the E_T^{30} solvent polarity value.

Compound	Solvent	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ε_{260}[d]	ε_{abs}[d]	ϕ[e]
	Dioxane	390	462	4.0	20.0	2.9	0.36
	0.90 - 0.10	384	482	5.4	19.0	2.6	0.45
	0.70 - 0.30	384	495	5.8	20.3	2.8	0.37
	0.30 - 0.70	373	514	7.4	19.5	2.5	0.14
	Water	365	526	8.4	17.0	2.5	0.03
	D_2O	365	527	8.4	16.4	2.6	0.10

53

Table 4.3. Photophysical data of ^{DMA}**C** (**53**) in solvent mixtures of dioxane and water. [a] λ_{abs} are reported at the most red-shifted absorbance wavelength in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1}. [d] Extinction coefficients (ε) are reported in 10^3 M^{-1} cm^{-1} and are given at λ_{abs} = 260 nm and at the most red-shifted absorbance. [e] Quinine hemisulfate (ϕ = 0.55 at λ_{ex} = 370 nm) in 0.5 M H_2SO_4 was used as the fluorescent standard for the relative quantum yields (ϕ). Errors associated are ±10% of the given values.

4.3.4. Protonation of ^{DMA}C

To mimic the nucleobase environment of DNA,[193] acetonitrile was used as a solvent for characterizing the protonation-dependent photophysical properties of ^{DMA}**C** (**Table 4.2**). Upon adding 1.0 equivalents of HCl, ^{DMA}**C** exhibited pronounced red-shifting of its absorbance (λ_{abs} = 395 → 425 nm) and emission wavelengths (λ_{em} = 485 → 555 nm, **Figure 4.8**). The addition of excess HCl resulted in dramatic blue-shifting (λ_{abs} = 425 → 325 nm, and λ_{em} = 555 → 425 nm). Taken together with our HOMO-LUMO calculations (**Figure 4.4**), these results suggest that *N*3 is the primary protonation site of ^{DMA}**C**. This conclusion is further supported by acid-mediated folding reactions of ^{DMA}**C**-modified DNA.

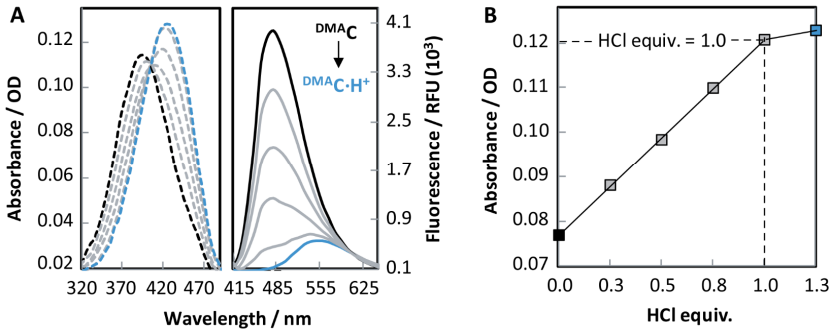


Figure 4.8. (A) Absorbance (---) and emission (—) spectra (λ_{ex} = 405 nm) of ^{DMA}**C** (black) and ^{DMA}**C**·H$^+$ (blue) at 40 μM in MeCN upon addition of HCl. (B) Plot between the absorbance intensity at λ_{abs} = 425 nm against the number of HCl equivalents added as a 4 M HCl stock solution in dioxane.

To measure the pK_a of protonated ^{DMA}**C** in water, its absorbance and emission spectra were evaluated as a function of pH (**Figure 4.9A**). Upon acidification, red-shifted absorbance maxima and a sigmoidal dependence of absorbance at 360 nm versus pH were observed that fit well to the Henderson-Hasselbalch equation for pK_a = 4.5 (**Figure 4.9B**). This is the same pK_a value as native cytidine.[31-32] Similar results were obtained with the fluorescence data (**Figure 4.9C**). The absorbance spectra of ^{DMA}**C** also revealed an emergence of absorbance at 440 – 480 nm with decreasing pH (**Figure 4.9A**). This suggested ^{DMA}**C** might be used as a "turn-on" fluorescent probe for *i*-motif formation when selectively excited at these wavelengths.

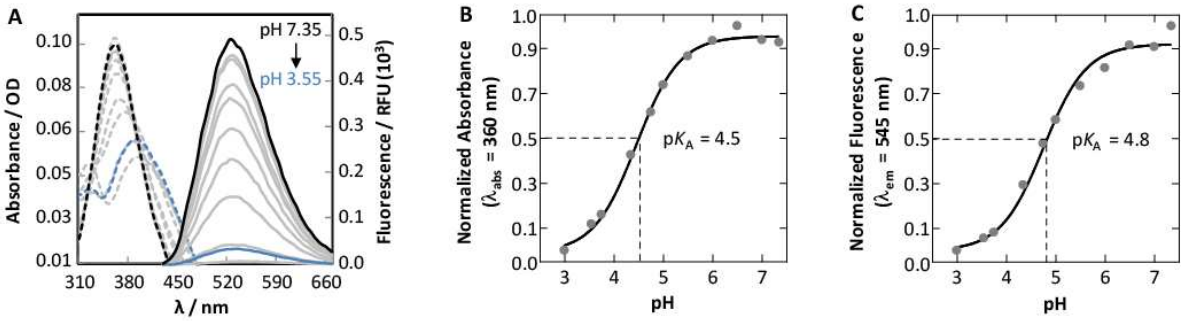


Figure 4.9. (A) Absorbance (---) and emission (—) spectra (λ_{ex} = 392 nm) of ^{DMA}**C** (black) and ^{DMA}**C**·H$^+$ (blue) at 40 μM in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl). (B) Plot of absorbance intensity at λ_{abs} = 360 nm versus pH. (C) Plot of fluorescence intensity at λ_{em} = 545 nm versus pH.

4.4. Shedding Light on DNA Folding

4.4.1. Synthesis of ^{DMA}C-Modified DNA

To facilitate site-specific incorporation of ^{DMA}**C** into DNA, synthesis of phosphoramidite **56** commenced with acetylation of nucleoside **52** to give **54** in 83% yield (**Figure 4.10**). Following silyl group removal, DMT protection of the 5'-hydroxyl group furnished **55** in a yield of 56% over two steps. Phosphitylation of the 3'-hydroxyl group gave **56** in an isolated yield of 71%. Phosphoramidite **56** was compatible with standard, automated DNA synthesis, although some deamination to give DNA containing the corresponding thymidine analog was observed as a minor side product. This could be removed by HPLC to give pure ^{DMA}**C**-modified oligonucleotides (**Figure 4.11** and **Table 7.2** in the Experimental Chapter).

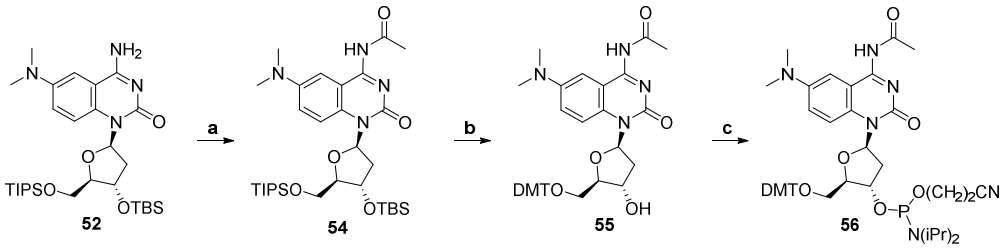


Figure 4.10. Synthesis of ^{DMA}**C** phosphoramidite **(56)** Reaction conditions: a) Ac_2O (1.2 equiv.), DMAP (10 mol%), pyridine, 23°C, 1 h, 83% yield. b) (i) TBAF, THF, 23 °C, 2 h, (ii) DMTCl (1.2 equiv.), pyridine, 23 °C, 45 min, 56% yield (over two steps). c) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (2.0 equiv.), DIPEA, CH_2Cl_2, 0 °C, 45 min, 71% yield.

The cytosine-rich human telomeric repeat sequence $(^{5'}CCCTAA^{3'})_n$ was synthesized to contain ^{DMA}**C** at four different positions. This sequence (where n = 4) folds into two closely related *i*-motif conformations *in vitro* at pH = 6.0, each one containing six C-C$^+$ base pairs and three TAA loops (**Figure 4.11**).[22,246-248] ^{DMA}**C** was incorporated at positions stacked between two other C-C$^+$ base pairs "**TeloX2** and **TeloX14**"; or at positions proximal to the loops "**TeloX13** and **TeloX15**" where **X** = ^{DMA}**C**.

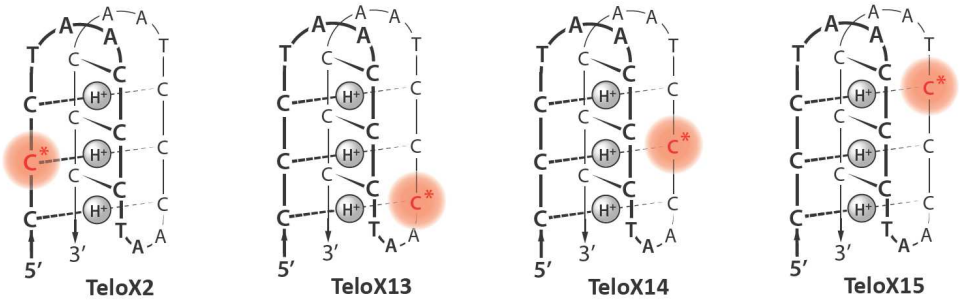


Figure 4.11. Secondary structure of a human telomeric *i*-motif containing the ^{DMA}**C** probe at four different positions. The dominant *i*-motif conformation "5'E" is shown.[22,246-248] See **Table 7.2** in the Experimental Chapter for all primary DNA sequences used in this study.

4.4.2. Impact of ^{DMA}C on DNA Structure and Stability

Circular dichroism (CD) and thermal denaturation (T_m) studies were used to evaluate the impact of ^{DMA}**C** on global structure and stability of duplex and *i*-motif DNA (**Figure 4.12** and **Table 4.4** and additional spectra in **Appendix A9 – A11**).

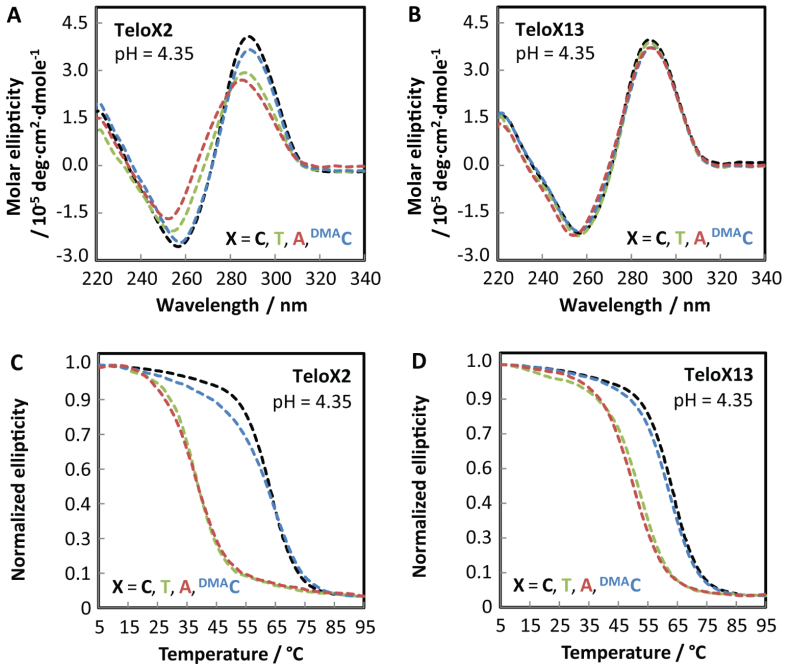


Figure 4.12. CD spectra at 25 °C of *i*-motif DNA of (A) **TeloX2** and (B) **TeloX13** with X = C (black), T (green), A (red) and ^{DMA}**C** (blue) at pH = 4.35. Thermal denaturation spectra of *i*-motif DNA of (A) **TeloX2** and (B) **TeloX13**. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) at pH = 4.35.

Oligonucleotides containing either **X** = C (wild type), ^{DMA}**C**, T or A at each site of incorporation (**X2**, **X13**, **X14**, and **X15**) were folded into *i*-motifs by heating and cooling single-stranded oligonucleotides in a phosphate citric acid buffer containing 200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl (pH = 4.35 or 5.50). The resulting CD spectra exhibited maxima at λ_{max} = 288 nm, minima at λ_{min} = 256 nm, and cooperative unfolding upon heating that are characteristic features of *i*-motif structures (**Figure 4.12**).[95] ^{DMA}**C** caused little, if any impact on the global structure or stability of *i*-motifs as compared to the wild type (wt) sequence containing cytosine (ΔT_m = +0.4 to -2.8 °C, **Table 4.4**). In contrast, mutant constructs containing A or T were dramatically destabilized (ΔT_m = -7.1 to -26.0 °C, **Table 4.4**). These results demonstrate that *i*-motif structures are exquisitely sensitive to base pair mismatches, and that ^{DMA}**C** is an excellent mimic of cytidine in this demanding context. This conclusion is further supported by pH-dependent *i*-motif folding experiments (**Figure 4.14**).

Sequence	**X**	T_m *i*-motif at pH = 4.35	T_m *i*-motif at pH = 5.50	T_m duplex at pH = 7.35
Telo(wt)	C	62.6	38.5	64.9
TeloX2	T	38.3 (-24.3)	27.6 (-10.9)	63.1 (-1.8)
″	A	37.8 (-24.5)	25.3 (-13.2)	61.1 (-3.8)
″	^{DMA}**C**	61.7 (-0.9)	38.9 (+0.4)	64.5 (-0.4)
TeloX13	T	51.2 (-11.4)	31.4 (-7.1)	57.3 (-7.6)
″	A	49.6 (-13.0)	29.6 (-8.9)	58.3 (-6.6)
″	^{DMA}**C**	61.2 (-1.4)	37.6 (-0.9)	64.1 (-0.8)
TeloX14	T	38.7 (-23.9)	28.1 (-10.4)	56.9 (-8.0)
″	A	36.6 (-26.0)	27.8 (-10.7)	58.8 (-6.1)
″	^{DMA}**C**	61.6 (-1.0)	38.7 (+0.2)	65.9 (+1.0)
TeloX15	T	46.7 (-15.0)	28.5 (-10.0)	56.7 (-8.2)
″	A	45.5 (-17.1)	27.5 (-11.0)	57.5 (-7.4)
″	^{DMA}**C**	59.8 (-2.8)	36.9 (-1.6)	66.3 (+1.4)

Table 4.4. Thermal denaturation melting temperatures (T_m). [a] Values are given as T_m (+/-ΔT_m) where ΔT_m is the deviation from the wild-type (wt) construct **TeloC**. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) at pH = 4.35, pH = 5.50 and pH = 7.35. Reproducibility is within 0.3 °C.

Duplex DNA structures were prepared by heating and cooling each oligonucleotide in the presence of 1.1 equivalents of complement strand. The CD spectra of the resulting duplexes exhibit characteristic maxima at λ_{max} = 264 nm, and minima at λ_{min} = 240 nm (**Appendix A11**).[95,261] ^{DMA}**C** caused little, if any perturbation of the global structure or thermal stability of duplex DNA as compared to the wild type (ΔT_m = +1.4 to -0.8 °C, **Table 4.4**). In contrast, mutant constructs containing A-G and T-G mismatches exhibited significant thermal destabilizations as compared to wild type (ΔT_m = -1.8 to -8.2 °C, **Table 4.4**). Taken together, these results demonstrate that ^{DMA}**C** is an excellent mimic of cytidine in both duplex and *i*-motif structures.

4.4.3. pH-Dependent DNA Folding by CD and Fluorescence

CD and fluorescence spectroscopy were used to characterize the pH-dependent folding of single-stranded oligonucleotides into *i*-motif DNA.[42,95,261] At pH ≥ 6.50, the wild type oligonucleotide exhibits a CD maximum at λ_{max} = 274 nm and minimum at λ_{min} = 249 nm, consistent with a disordered oligonucleotide (**Figure 4.14A**).[42,95,261] When wild-type samples were heated and cooled at variable pH's, a transition from unfolded to *i*-motif structures was observed with a ½ effective pH value (pH_m) of 6.0 (**Figure 4.14B**). ^{DMA}**C**-modified oligonucleotides also gave pH_m values = 6.0 ± 0.2 at all four sites of incorporation according to pH-dependent changes in fluorescence (**Figure 4.14D**, **E** and **F**). By selective excitation of protonated ^{DMA}**C** at 480 nm, a 5-fold increase in fluorescence emission was observed upon *i*-motif folding, with a pH_m value = 5.9 (**Figure 4.13C** and **D**). As for native cytosine residues, this value is 1.4 p*K* units higher than the pK_a value of ^{DMA}**C** in bulk solution (**Figure 4.9**).

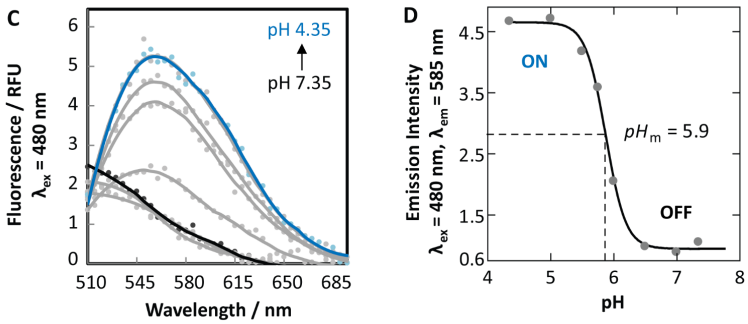


Figure 4.13. (C) Fluorescence spectra (λ_{ex} = 480 nm) of **TeloX13** (**X** = ^{DMA}**C**) at pH = 4.35 – 7.35. (D) Plot of fluorescence intensity (λ_{em} = 585 nm) versus pH.

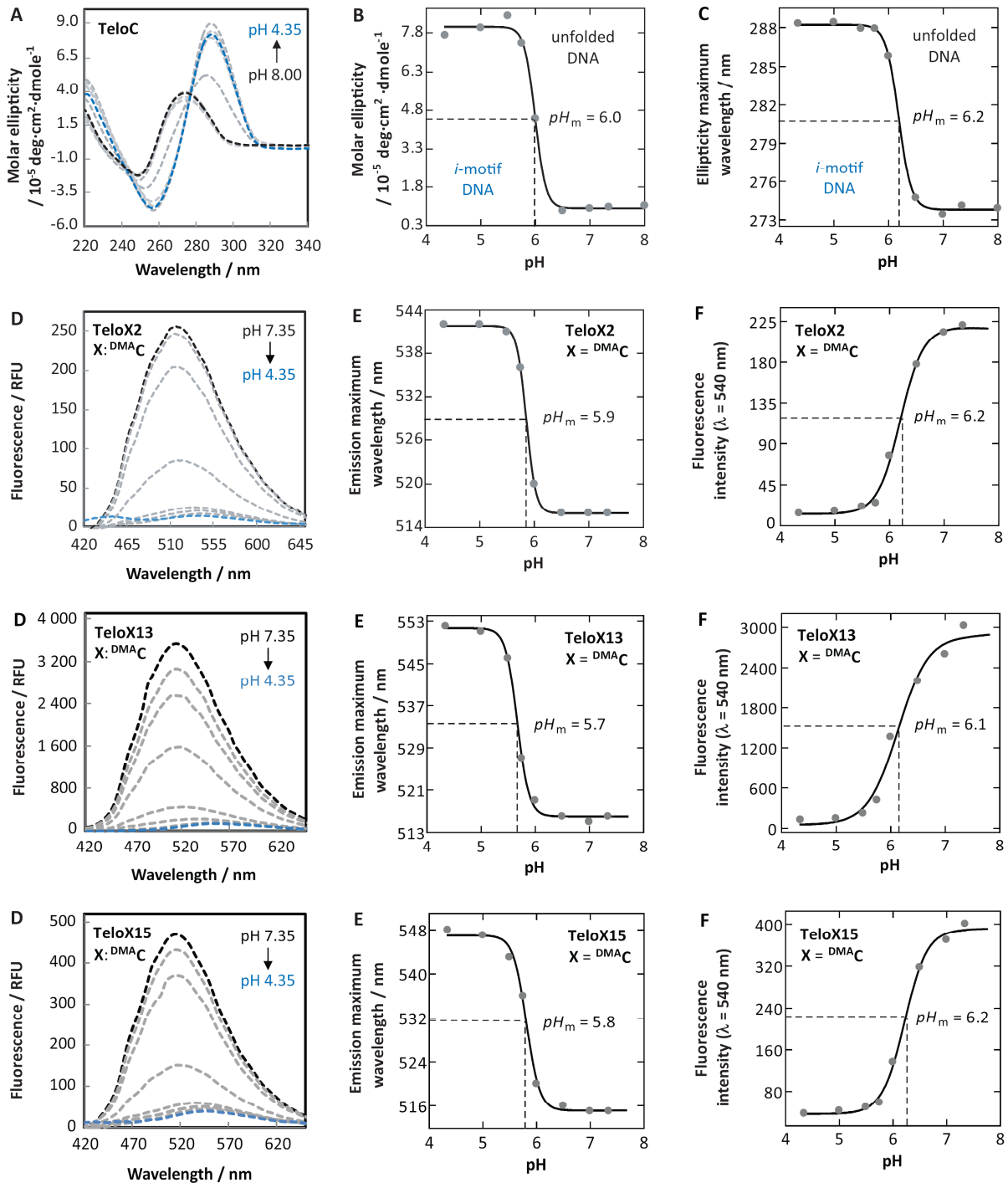


Figure 4.14. (A) CD spectra at 25 °C at pH = 4.35 (blue), 8.00 (black) and intermediate values (grey) in buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl). (B) Plot of molar ellipticity (λ_{abs} = 292 nm) and (C) plot of ellipticity maximum wavelength versus pH. (D) Fluorescence spectra (λ_{ex} = 260 nm) at pH = 4.35 – 7.35. Plot of (E) emission maximum wavelength and (F) fluorescence intensity (λ_{em} = 540 nm) versus pH.

4.4.4. Fluorescence Properties of ^{DMA}C in DNA

The fluorescence properties of ^{DMA}**C** are highly sensitive to DNA conformation (**Table 4.5**). ^{DMA}**C** in the context of *i*-motif structures exhibits ca. 10-fold decreased quantum yields, and ca. 35 nm red-shifted absorption (λ_{abs} = 400 – 407 nm) and emission maxima (λ_{em} = 542 – 551 nm, **Table 4.5**) as compared to unfolded DNA. These trends, upon DNA folding, resemble those observed for the ^{DMA}**C** nucleoside (**53**) upon its protonation in acetonitrile (λ_{abs} = 425 nm, λ_{em} = 555 nm, **Figure 4.8**).

Structure	Sequence	λ_{abs}[a]	λ_{em}[b]	ϕ[c]	ηt[d]	r[e]
unfolded	**TeloX2**	365	517	0.04	0.08	0.04
	TeloX13	378	516	0.09	0.11	0.04
	TeloX14	365	520	0.06	0.06	0.04
	TeloX15	370	516	0.03	0.12	0.05
i-motif	**TeloX2**	400	542	0.004	0.12	0.24
	TeloX13	405	551	0.01	0.15	0.17
	TeloX14	404	545	0.005	0.13	0.27
	TeloX15	407	548	0.006	0.17	0.21
duplex	**TeloX2**	365	512	0.02	0.05	0.15
	TeloX13	365	505	0.11	0.04	0.14
	TeloX14	365	505	0.06	0.03	0.17
	TeloX15	365	502	0.04	0.08	0.16

Table 4.5. Photophysical data of ^{DMA}**C** in DNA (**X** = ^{DMA}**C**) [a] λ_{abs} are reported for the most red-shifted absorbance maxima, [b] λ_{em} in nm. [c] ^{DMA}**C** (ϕ = 0.03 at pH = 7.35 with λ_{ex} = 370 nm and ϕ = 0.04 at pH = 5.50 with λ_{ex} = 410 nm) was used as the fluorescent standard for the relative quantum yields (ϕ) of ^{DMA}**C** in DNA, [d] DNA to probe energy transfer efficiencies (ηt) were calculated at λ_{ex} = 260 nm and at the maximum λ_{em} (See the Experimental Chapter for calculation details), [e] fluorescence anisotropy (r) calculated at pH = 7.35 (λ_{ex} = 370 nm, λ_{em} = 535 nm) and at pH = 5.50 (λ_{ex} = 410 nm, λ_{em} = 550 nm). All measurements were performed with 4 μM of DNA in phosphate citric acid buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl).

Fluorescence anisotropy (r) provides a powerful means to evaluate the dynamic motions of fluorophores.[108] Much larger anisotropy values were observed for ^{DMA}**C** in the context of *i*-motif structures (r = 0.17 – 0.27) as compared to unfolded oligonucleotides (r = 0.04 – 0.05, **Table 4.5**). Intermediate values were observed for duplexes (r = 0.14 – 0.17). Given the larger hydrodynamic radii of duplex and unfolded DNA as compared to *i*-motif,[262] these results suggest that local nucleobase dynamic

motions are highly restricted in C-C$^+$ base pairs. This conclusion is supported by NMR proton-deuteron exchange experiments demonstrating that the protons in "internal" C-C$^+$ base pairs exchange very slowly (lifetime ≈ 1 hour) with bulk solvent,[25] whereas exchange rates at sites neighboring loops are much faster (≈ 1 ms).[25] Consistent with this, the anisotropy values of ^{DMA}**C** were lower (r = 0.17 – 0.21) in C-C$^+$ base pairs proximal to loops (**TeloX13** and **TeloX15**), as compared to "internal" C-C$^+$ base pairs of the i-motifs (r = 0.24 – 0.27, **TeloX2** and **TeloX14**, **Table 4.5**).

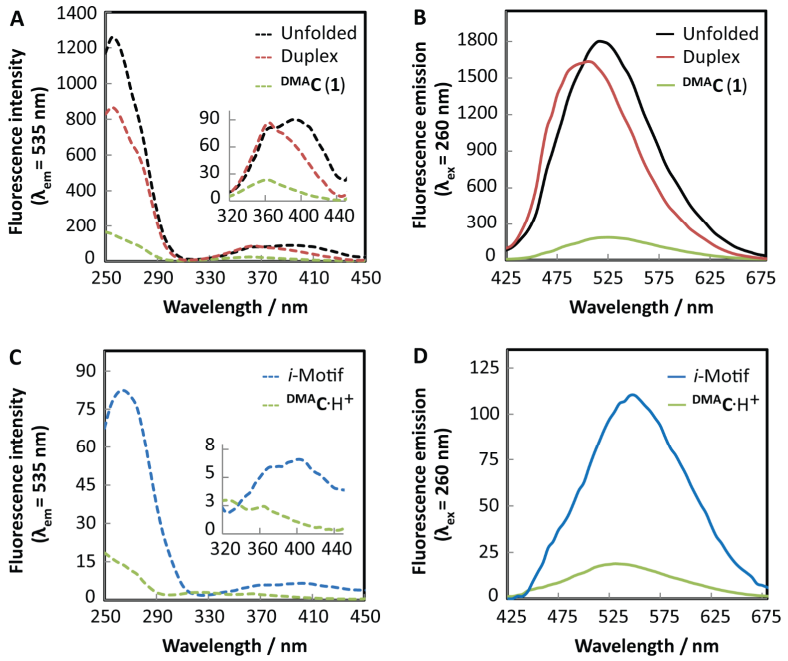


Figure 4.15. (A) Excitation (λ_{em} = 535 nm) and (B) emission (λ_{ex} = 260 nm) spectra of **TeloX13** at pH = 7.35 (**X** = ^{DMA}**C**). (C) Excitation and (D) emission spectra at pH = 5.50. DNAs were prepared as unfolded (black), duplex (red) and *i*-motif structures (blue). The spectra of the ^{DMA}**C** nucleoside (pH = 7.35) and ^{DMA}**C-H**$^+$ (pH = 3.55) are shown in green. All samples contained 4 μM of ^{DMA}**C**-modified DNA or ^{DMA}**C** nucleoside (**1**) in phosphate citric acid buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl).[263]

The excitation spectra of ^{DMA}**C**-containing oligonucleotides exhibited two broad maxima centered at λ_{ex} ≈ 260 and 375 nm (**Figure 4.15** and **Appendix A12**). The intensity at 260 versus 370 nm was much higher for the ^{DMA}**C**-containing oligonucleotides as compared to the ^{DMA}**C** nucleoside (**53**). This revealed the presence of energy transfer from unmodified bases (λ_{em} = 327 nm)[264] to ^{DMA}**C**.[265-267] Energy transfer efficiencies (η_t) were much higher in *i*-motif structures (η_t = 0.17 – 0.12) as compared to the same oligonucleotides prepared as duplexes (η_t = 0.08 – 0.03, **Table 4.5**). The exceptionally high nucleobase-

to-nucleobase energy transfer efficiencies in *i*-motif structures are consistent with X-ray crystallographic analyses demonstrating that *i*-motifs have high nucleobase densities and C-C^+ base pairs with unusually close base-stacking distances (3.1 Å) as compared to duplex DNA (3.4 Å).[26-27]

4.4.5. Assignment of Static Structure Using ^{DMA}C

The photophysical properties of ^{DMA}**C** provide well resolved and characteristic information about the folded structure of the DNA containing it. At all four sites of incorporation, ^{DMA}**C** emission maxima exhibited the following trends: *i*-motif (λ_{em} = 542 – 552 nm) > unfolded (λ_{em} = 516 – 520 nm) > duplex (λ_{em} = 502 – 512 nm); fluorescence anisotropy values: *i*-motif (r = 0.17 – 0.27) > duplex (r = 0.14 – 0.17) > unfolded (r = 0.04 – 0.05); energy transfer efficiencies: *i*-motif (η_t = 0.12 – 0.17) > unfolded (η_t = 0.06 – 0.12) > duplex (η_t = 0.03 – 0.08, **Table 4.5** and **Figure 4.15**). By measuring multiple photophysical parameters in parallel, ^{DMA}**C** can thereby be used to unambiguously assign the folded state of the DNA. For example, fluorescence anisotropy can easily differentiate unfolded (r = 0.04 – 0.05) versus folded DNA (r = 0.14 – 0.27), and fluorescence emission wavelengths can differentiate *i*-motif (λ_{em} = 542 – 552 nm) from duplex structures (λ_{em} = 502 – 512 nm).

4.4.6. Probing Real-Time DNA Dynamics with ^{DMA}C

To evaluate the ability of ^{DMA}**C** to serve as a real-time probe of DNA dynamics, strand displacement assays were conducted. In this approach, ^{DMA}**C**-containing oligonucleotides were pre-hybridized with a complementary sequence containing a variable-length "toehold" (TH) region to give "**GC**" duplexes containing a short, single-stranded overhang (**Figure 4.16** and **Table 7.5** in the Experimental Chapter). Displacement of the labeled strand is accomplished by adding a 4-fold excess of an unlabeled invading strand "**I**" to give a more stable duplex **"GI"** containing no single-stranded overhang.

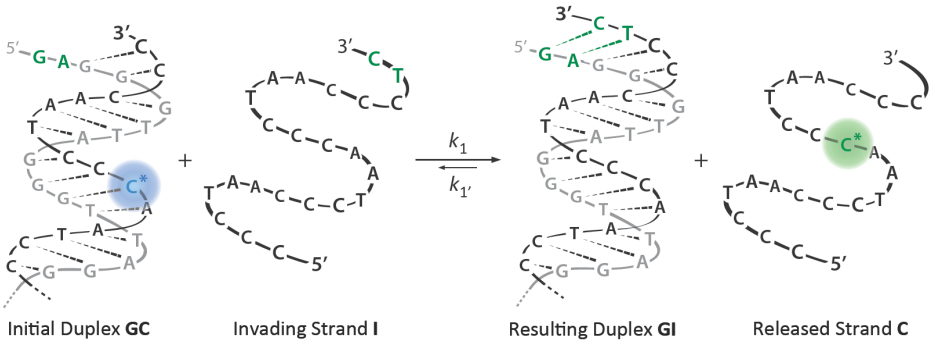


Figure 4.16. Representation of strand displacement reaction at pH = 7.35 with **TeloX13** (**X** = ^{DMA}**C**) and a toehold of 2 nucleotides (green).

The displacement of ^{DMA}**C**-containing strands was monitored in real time using fluorescence anisotropy, and rate constants (k) were calculated using pseudo-first approximations as previously described (**Figure 4.17** and **Appendix A13**).[268-270] The exact site of ^{DMA}**C** incorporation had very little ($\leq$ 2-fold) impact on the measured k values (**Table 4.6**), but large variations were observed with increasing toehold (TH) lengths: $k_{TH1} = 1.1 \times 10^1 < k_{TH2} = 1.8 \times 10^1 < k_{TH3} = 2.1 \times 10^2 < k_{TH4} = 4.6 \times 10^2$ $M^{-1} s^{-1}$ (**Table 4.6**). These rate constants were very similar to those previously reported using FRET-based constructs,[271] demonstrating that a single ^{DMA}**C** residue can replace the bulky end-labels used in FRET-based assays. These results further demonstrate that the rate limiting step of strand displacement is initiation rather than propagation, and that ^{DMA}**C** does not pose a significant barrier to either process.

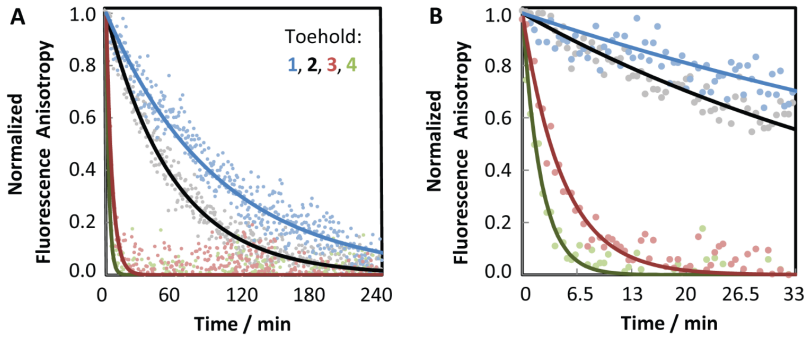


Figure 4.17. Real-time strand displacement according to fluorescence anisotropy of ^{DMA}**C**-modified oligonucleotide **TeloX13** (**X** = ^{DMA}**C**, λ_{ex} = 370 nm and λ_{em} = 535 nm). The lines represent the mono-exponential fit of the data to eq. 4.

Sequence	TH = 1	TH = 2	TH = 3	TH = 4
TeloX2	1.8×10^1	3.4×10^1	3.1×10^2	6.1×10^2
TeloX13	1.1×10^1	1.8×10^1	2.1×10^2	4.6×10^2
TeloX14	1.0×10^1	1.8×10^1	1.7×10^2	4.2×10^2
TeloX15	0.9×10^1	1.3×10^1	1.7×10^2	2.7×10^2

Table 4.6. Strand displacement rate constants (k) in $M^{-1}\,s^{-1}$ at pH = 7.35 for different "toehold" (TH) lengths where **X** = ^{DMA}**C.** Rate constants k_1 were calculated by fitting fluorescence anisotropy data (λ_{ex} = 370 nm, λ_{em} = 535 nm) to a single exponential decay eq. 4 (See Experimental Chapter). Measurements were performed using 4 μM of DNA in phosphate citric acid (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) buffer at pH = 7.35. The oligonucleotide sequences used are given in **Table 7.5** (Experimental Chapter).

4.4.7. i-Motif Structures Pose Large Kinetic Barriers to Duplex Formation.

To investigate the kinetic stability of *i*-motif structures, the invading "**I**" strand was mixed with ^{DMA}**C**-containing "**GC**" duplexes under slightly acidic conditions (**Figure 4.18**). Given the negligible impact that ^{DMA}**C** has on *i*-motif stability, the net thermodynamic driving force for *i*-motif unfolding, strand displacement, and *i*-motif refolding of the released strand "**C**" is approximately the same as in the "simple" stand-displacement experiments (**Figure 4.16**). pH-jump studies previously demonstrated that unimolecular *i*-motif folding of single-stranded DNA occurs very rapidly ($t_{½} \approx 1$ s) at pH = 5.7, while unfolding is much slower ($t_{½} \approx 100$ s) at pH = 6.5.[272] We therefore anticipated that the rate-limiting step of strand displacement (**Figure 4.18**) would be unfolding of the "**I**" strand prior to its invasion of the "**GC**" duplex.

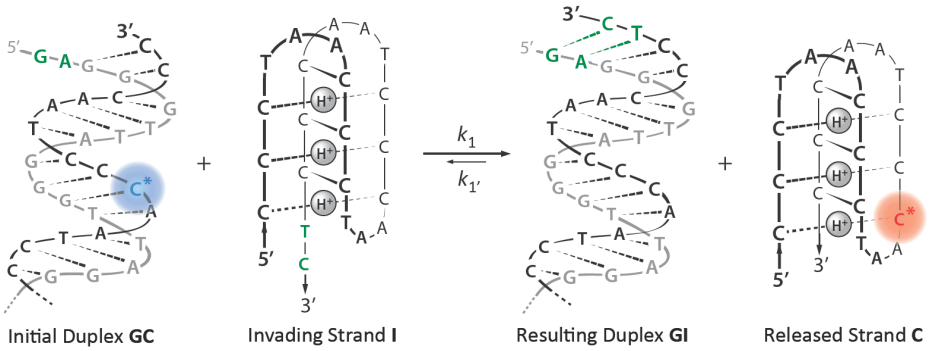


Figure 4.18. Representation of strand displacement reaction at pH ≤ 6.50 with **TeloX13** (**X** = ^{DMA}**C**) and a toehold of 2 nucleotides (green).

^{DMA}**C**-modified oligonucleotides exhibit large differences in fluorescence intensity in duplex versus *i*-motif structures (**Figure 4.15** and **Appendix A12**). We therefore monitored the fluorescence intensities of ^{DMA}**C**-containing "GC" duplexes upon addition of a 4-fold excess of invading stands "**I**" at various pH values (**Figure 4.19**). At pH = 7.35, where the **I** strand is unstructured, nearly identical rate constants were obtained using both fluorescence anisotropy (**Table 4.6**) and fluorescence intensity assays (**Table 4.7**). According to CD, the **I** strand exists as a 1:1 mixture of *i*-motif folded and unfolded DNA at pH = 6.0 (**Figure 4.14A**), yet 10-fold lower strand displacement rates were observed at pH = 6.0 as compared to pH = 7.35 (**Figure 4.19** and **Table 4.7**). This supports the fact that *i*-motif folding is not a simple two-state process,[272-273] and that all "**I**" strands possess at least some partial *i*-motif character at pH = 6.0 that impacts the rate of strand invasion. According to CD, the invading strand "**I**" is fully folded into an *i*-motif at pH = 5.75 (**Figure 4.14A** and **Figure 4.21**) that exhibits only marginal thermal stability (T_m = 31 – 38 °C, **Figure 4.21**). The pre-folded "**GC**" duplex, in contrast, exhibits high kinetic and thermodynamically stability over the entire pH range of 5.75 – 7.35 (T_m = 61 – 65 °C, **Figure 4.22**).[274] Upon mixing the two structures at pH = 5.75, strand displacement kinetics were much slower than those observed at pH = 6.0 (**Figure 4.19** and **Table 4.7**). In stark contrast to the results from the "simple" strand displacement experiments (**Table 4.6**), there was almost no difference in rate constants obtained for toehold (TH) = 2 versus TH = 3 at pH = 5.75 (**Table 4.7**). These results validate our premise that *i*-motif unfolding is the rate limiting step of strand displacement at pH = 5.75.

Sequence	pH	TH = 2	TH = 3
I	7.35	4.1×10^1 (0.90)	3.2×10^2 (0.73)
TeloX13	6.50	2.9×10^1 (0.87)	1.7×10^2 (0.77)
X = ^{DMA}C	6.00	0.54×10^1 (0.55)	4.1×10^1 (0.68)
	5.75	0.15×10^1 (0.66)	0.10×10^1 (0.68)

Table 4.7. Strand displacement rate constants (k_1) in M^{-1} s^{-1} for different "toehold" (TH) lengths. Rate constants k_1 and k_2 with **I** were calculated from a bi-exponential fit of fluorescence intensity data to eq. 5 (See Experimental Chapter). The dominant component (k_1) of each fit is reported, and its relative contribution as compared to k_2 is given in parentheses. The minor components (k_2) were all pH-independent, with values approximately equal to 1.8×10^1 M^{-1} s^{-1}. Similar values were measured by the "simple" strand displacement assay for TH = 2, suggesting that k_2 reflects strand displacement by single-stranded **I**. All measurements were performed using 4 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl).

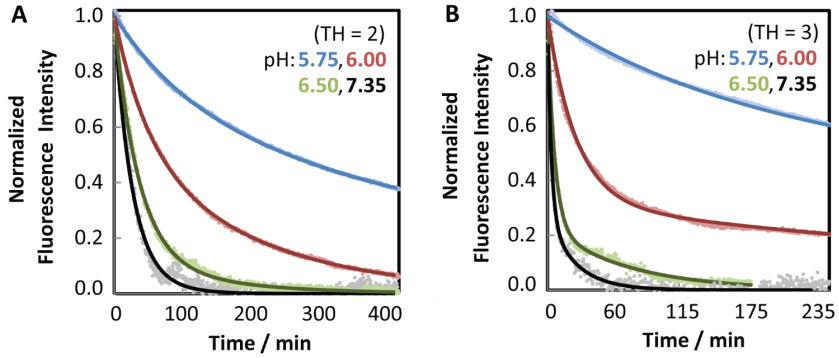


Figure 4.19. Strand displacement kinetics at pH from 7.35 to 5.75 monitored by the fluorescent intensity of **TeloX13** (**X** = ^{DMA}**C**) with toehold (A) TH = 2 and (B) TH = 3 (λ_{ex} = 370 nm and λ_{em} = 535 nm). Similar results were obtained for ^{DMA}**C** incorporated at **TeloX2**. The points represent the measured fluorescence data and the lines represent the bi-exponential fitting to eq. 5.

To evaluate the possibility that the small differences in pH were responsible for the decreased rate constants by an i-motif-independent mechanism, a closely-related invading strand mutant "$\mathbf{I_{mut}}$" was also evaluated (**Figure 4.21** and **Table 7.6** in the Experimental Chapter). $\mathbf{I_{mut}}$ contains a central T-T base pair mismatch and is therefore incapable of forming a stable *i*-motif at pH = 5.75 and 25 °C. Upon mixing ^{DMA}**C**-containing "GC" duplexes with $\mathbf{I_{mut}}$ oligonucleotides, little or no changes in rate constants were observed over the entire pH range 5.5 – 7.35, with values approximately equal to k_{TH2} = 1.2 × 10^2 M^{-1} s^{-1} and k_{TH3} = 5.8 × 10^2 M^{-1} s^{-1} (**Table 4.8** and **Figure 4.20**). Taken together, these results prove that the pH-dependent effects exhibited by the wild type "**I**" are due to *i*-motif unfolding. A comparison of the rates measured at pH = 7.35 versus 5.75 reveals that *i*-motif unfolding presents large kinetic barrier (320-fold difference for TH = 3) to duplex formation under near-physiological conditions of constant pH, temperature and salt.

Sequence	pH	TH = 2	TH = 3
$\mathbf{I_{mut}}$	7.35	1.7 × 10^2	5.5 × 10^2
TeloX13	6.50	1.0 × 10^2	7.5 × 10^2
X = ^{DMA}**C**	6.00	0.9 × 10^2	7.6 × 10^2
	5.75	1.1 × 10^2	2.4 × 10^2

Table 4.8. Strand displacement rate constants (k_1) in M^{-1} s^{-1} for different "toehold" (TH) lengths. Rate constants *k* with $\mathbf{I_{mut}}$ were calculated by fitting the fluorescent intensity of **TeloX13** (λ_{ex} = 370 nm, λ_{em} = 515 nm) to a mono exponential decay eq. 4 (See Experimental Chapter). All measurements were performed using 4 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl).

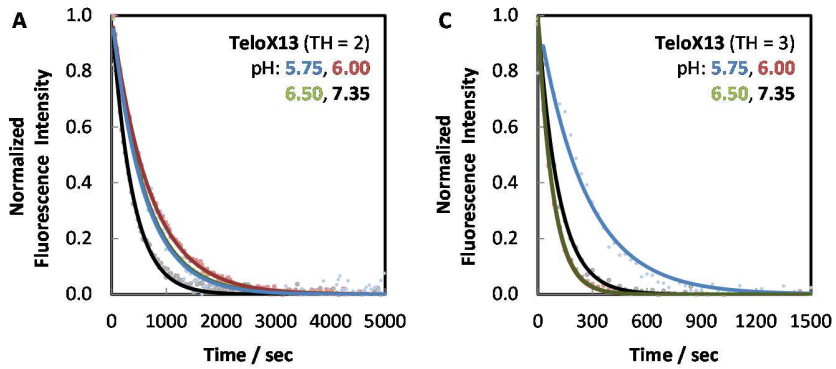


Figure 4.20. Strand displacement kinetics at pH ranging from 7.35 to 5.75 monitored by the fluorescent intensity of **TeloX13** (**X** = ^{DMA}**C**) with toehold (A) TH = 2 and (C) TH = 3 (λ_{ex} = 370 nm and λ_{em} = 515 nm). The points represent the measured fluorescence data and the lines represent the mono-exponential fitting to eq. 4.

4.4.8. Secondary Structures of DNAs Used in Strand Displacement Assays

Circular dichroism and thermal denaturation spectra were used to determine the secondary structure of the oligonucleotides used in the kinetic experiments at pH = 5.75 and 25 °C (**Figure 4.21**). The following data demonstrate that the wild-type (**X** = **C**) invading oligonucleotides (**I**) are thermodynamically folded into an *i*-motif structure (λ_{abs} = 288 nm). The mismatched (**X** = **T**) invading oligonucleotides ($\mathbf{I_{mut}}$) are unstructured in this condition (λ_{abs} = 274 nm).

CD titrations were performed to determine the thermodynamic equilibrium between *i*-motif and duplex DNA as a function of pH (**Figure 4.22**). At pH < 4.0, only the *i*-motif structure is observed, whereas at pH > 5.0, only the duplex DNA exists. As expected, the pH_m observed is similar to the pK_A of *N*3 protonated cytosine (pH_m = pK_A = 4.5), which indicates that the protonation of the cytosines is the driving force for the folding of the *i*-motif structure. This demonstrates that a stable initial duplex (**GC**) is formed at pH = 5.75.

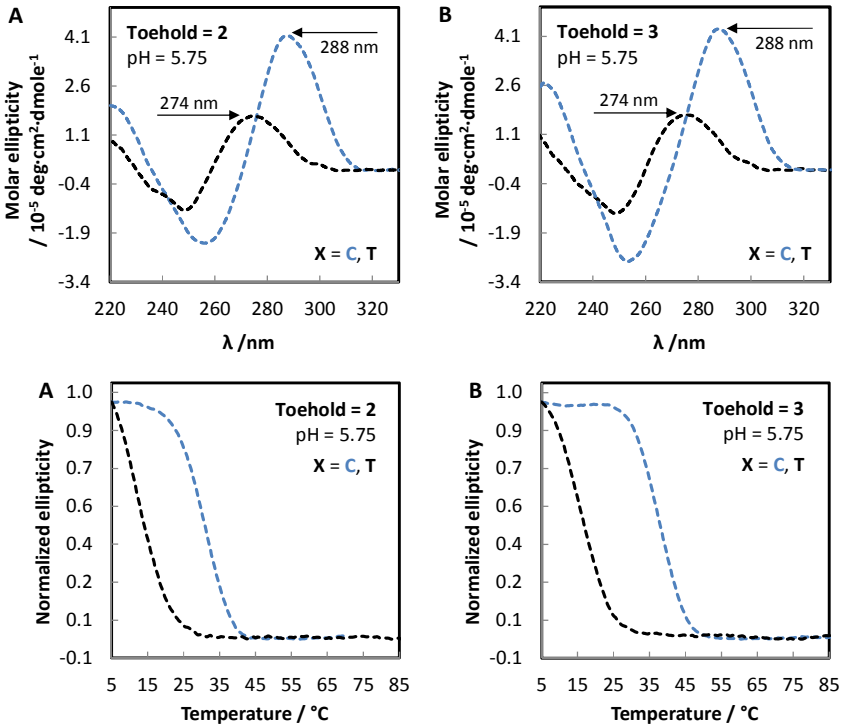


Figure 4.21. CD spectra at 25 °C and pH = 5.75 of the invading oligonucleotides (**I**) with toehold of (A) TH = 2 and (B) TH = 3 with **X** = **C** and **T**. Thermal denaturation spectra of the invading oligonucleotides (**I**) with toehold of (A) TH = 2 and (B) TH = 3 with **X** = **C** and **T**. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl).

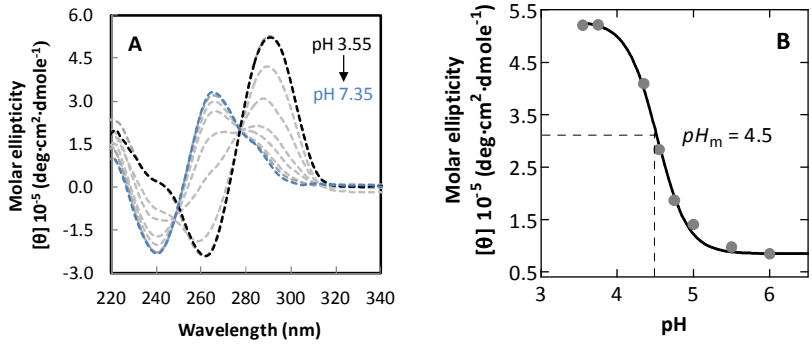


Figure 4.22. (A) CD spectra at 25 °C of wild-type duplex DNA and *i*-motif upon titration at pH ranging from 3.25 to 7.00. (B) Plot between the molar ellipticity at λ = 292 nm and the pH. All samples contain 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl at pH = 4.35 to 7.35). Double-stranded DNA contains 1.0 equivalents of the complementary strand

4.5. Discussion and Conclusions

Due to their relevance in biological and materials sciences, interest in DNA structures containing non-canonical base pairs is rapidly growing. Over the past decade, thermodynamically-stable structures such as G-quadruplexes have received much attention,[11,15,121,245,275] whereas relatively little focus has been paid to their complementary DNA sequences that can fold into *i*-motif structures.[34,38] One reason for this bias is that *i*-motif structures exhibit little or no thermodynamic stability at pH $\geq$ 7.2 in the absence of stabilizing factors.[35-36,243] The largest and most stable *i*-motif structures reported to date exhibit thermal denaturation temperatures of T_m = 27 – 32 °C at pH 7.0.[34] Despite these low thermodynamic stabilities, mounting evidence suggests that *i*-motif structures may indeed possess biological functions. For example, a variety of natural proteins are known to selectively bind to and unfold *i*-motif structures.[48-49] As such, molecules that stabilize *i*-motif structures can potentially inhibit endogenous DNA unfolding pathways. Consistent with this type of activity, *i*-motif-stabilizing ligands have been shown to inhibit telomerase activity in cells,[276] and can trigger a transcriptional switch between two alternative folding pathways involving competition between hairpin and *i*-motif conformations.[50-51] In the absence of stabilizing ligands, however, it is unclear how endogenous *i*-motif structures, with such marginal thermodynamic stabilities, could impact gene regulation and/or chromosome stability. One possibility is that *i*-motif structures could impose kinetic, rather than thermodynamic barriers to folding pathways involving duplex formation.

Previous studies aimed at characterizing the kinetic parameters of *i*-motif/single-strand DNA transitions have utilized variable temperature and/or pH-jump experiments to trigger folding/unfolding.[22,272] These studies demonstrated that unimolecular *i*-motif folding occurs very rapidly ($t_{½} \approx 1$ s) at pH = 5.7, while unfolding is much slower ($t_{½} \approx 100$ s) at pH = 6.5.[272] While this approach provides a direct means to study folding kinetics, pH/temperature jumps serve to initiate folding or unfolding reactions by perturbing the physical-chemical properties of entire system. The resulting data can therefore not be interpolated to constant conditions of pH and temperature, where competition between alternative conformations would normally take place. In addition, C-rich DNA sequences, with few exceptions,[277-278] are found in the context of double-stranded DNA, and therefore exist in complex equilibria involving duplex, G-quadruplex and *i*-motif structures.[59] Most biologically relevant pathways are therefore expected to involve competition between inter- versus intra-molecular folding pathways.[279]

To our knowledge, only a single previous study has attempted to quantify the kinetic barriers of duplex formation imposed by *i*-motif unfolding.[244] Sugimoto and co-workers used CD to evaluate duplex strand annealing rates upon mixing an *i*-motif with an "unfolded" single-stranded DNA under conditions of constant pH and temperature. Strand annealing at pH = 5.5 was approximately 10-fold slower than that measured at pH = 7.0. The authors concluded that *i*-motif unfolding was responsible, but residual DNA structure in the single-stranded G-rich DNA might have caused an underestimation of this effect.[280] Such residual structure is especially problematic in the case of G-rich strands, as it is extremely difficult in practice to eliminate all residual G-quadruplex-type structure.[281] These effects could potentially account for the relatively small, 10-fold differences in folding rates reported by this study, where the unfolding of the G-rich strands may have contributed to the rates measured at both pH values.

Strand displacement reactions are biologically relevant processes,[282-283] that are compatible with the study of DNA (re)folding reactions under constant conditions of temperature, pH and salt. We have developed a novel strand displacement assay where the addition of excess unlabeled DNA is used to initiate a thermodynamically favorable unfolding-refolding reaction that displaces a ^{DMA}**C**-modified oligonucleotide from duplex DNA. This approach eliminates the problems associated with using single-stranded G-rich (G-quadruplex) DNA for strand annealing.[281] Given the negligible impact that ^{DMA}**C** has on DNA stability, the net thermodynamic driving force of *i*-motif unfolding, strand displacement, and refolding of the released "**C**" strand is approximately the same for the addition of an unstructured invading strand "**I**" (**Figure 4.16**) as compared to an *i*-motif- folded "**I**" strand (**Figure 4.18**). At pH = 7.35, where the **I** strand is mostly unstructured, the kinetics of strand displacement are highly dependent on "toehold" (TH) length (**Table 4.6**). In contrast, when the **I** strand is folded into an *i*-motif structure at pH = 5.75, the overall kinetics of strand displacement are independent of the TH length (**Table 4.7**). These results indicate that *i*-motif unfolding is the rate-limiting step for duplex formation under slightly acidic conditions. The rate constants for strand displacement by *i*-motif DNA (k = 1.0 $M^{-1}s^{-1}$), were 320-fold lower than those measured for unfolded DNA (k = 3.2 × 10^2 $M^{-1}s^{-1}$). These results are surprising, given the low thermodynamic stabilities of *i*-motif structures derived from the human telomeric repeat (T_m ≈ 38 °C at pH = 5.75). These results reveal that nucleic acid structures with only marginal thermodynamic stabilities can still pose large kinetic barriers ($t_{1/2}$ > 12 hours) to duplex formation. These results further suggest that the high mechanical stability of *i*-motif structures measured using optical tweezers is likely to be a kinetic rather than a thermodynamic effect.[284] Taken together, these results suggest that energy-

dependent processes would be needed to resolve *i*-motif structures formed *in vivo*, especially in the context of acid-loving bacteria where intracellular pH values can be found in the range of 4.6 – 6.5.[230] The high kinetic stability of *i*-motif structures may provide a biological driving force for the evolution of *i*-motif-selective helicases and single-stranded DNA binding proteins that unfold *i*-motif structures.[48-49] The inhibition of *i*-motif unfolding could potentially lead to misregulation of certain biological processes *in vivo*, such as telomere capping,[276] and gene regulation.[50-51] Given the large kinetic barriers that *i*-motif structures can impose on strand displacement and duplex formation, small molecules that selectively inhibit the initiation and/or propagation of strand displacement could offer a novel therapeutic means to impact chromatin structure, function and dynamics.

Chapter 5 | Sensing Mismatched Base Pairs in Duplex DNA

Chapter 5

Sensing Mismatched Base Pairs in Duplex DNA

5.1. Summary

Human genetic variations are markers associated with complex diseases. The identification of these markers can lead to the development of customized medicine. To accurately report the most abundant genetic variations, known as single nucleotide polymorphisms (SNPs), a fluorescent nucleoside mimic (^{DMA}**C**) was introduced in the telomeric DNA sequence. ^{DMA}**C**'s photophysical properties including fluorescence intensity and anisotropy (r) were used to discriminate Watson-Crick base pairing with G ($r = 0.15 \pm 0.01$) in duplex DNA versus mismatch pairing with C, T or A ($r = 0.10 \pm 0.01$). ^{DMA}**C** can therefore serve as a base-discriminating fluorescent nucleoside analog for SNP typing.

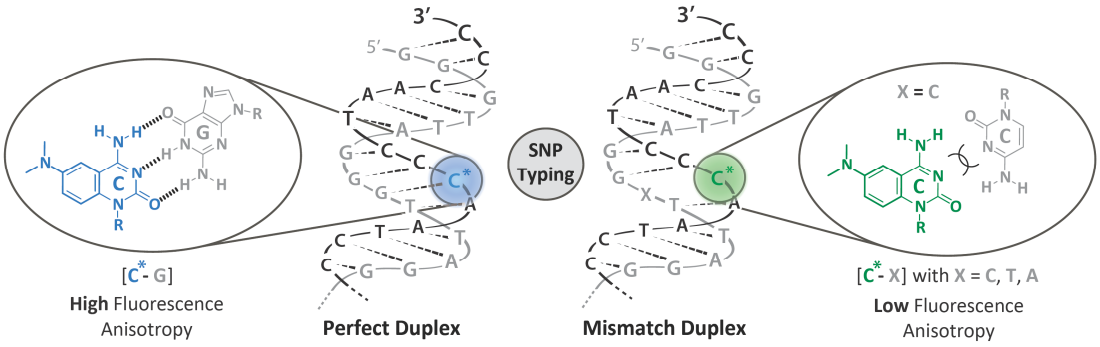


Figure 5.1. A fluorescent cytosine mimic (^{DMA}**C**) for sensing mismatched base pairing.

5.2. Introduction

5.2.1. Single-Nucleotide Polymorphisms

Single-nucleotide polymorphisms (SNPs) represent DNA sequences in which a single base pair mutation occurs to create a unique DNA pattern for each individual.[285-286] SNPs are the most abundant genetic variation and occur throughout the entire human genome at a frequency of approximately 0.1%.[287] The accumulated effects of SNPs are thought to result in larger variation in each individual's phenotype and may play important roles in diseases by affecting gene function.[288] For example, these mutated base pairs have been shown to cause allele-specific effects on the level of gene expression.[289] SNPs can therefore serve as markers for cancer,[290] and cardiovascular diseases,[291] to name only two.

Individuals can respond differently to the same drug and the identification of SNPs can be used to develop more effective therapies. When an SNP site is located in a coding region, the resulting protein may contain an alternative amino acid. This can then altered the binding receptors of a particular drug as well as enzymes controlling its metabolism. Characterizing SNPs in a selected population can lead to the development of customized medicine.[292] Because of the potentially important applications of SNPs in preventing and combating human diseases, typing methods to detect SNP's have become a growing research area.[293-294]

5.2.2. Detection of SNPs

Several approaches have been developed and commercialized to detect single base mismatches in DNA.[294] Most methods utilized a high fidelity polymerase in primer extension, but non-enzymatic techniques, such as FRET-based strand exchange can provide attractive alternatives. Recently, the development of base-discriminating fluorescent (BDF) nucleoside analogs has emerged as a powerful tool for SNP typing.[295] BDF probes are emissive nucleobases that can clearly distinguish matched versus mismatched base pairing according to changes in their photophysical properties.[296-297] This technique can provide a simpler and faster readout than conventional methods. The development of sensitive fluorescent nucleoside analogs with distinct photophysical properties in the context of mismatch duplex is therefore of broad interest. The design of such tools remains, however, a complicated task, since the relationships between chemical structure and photophysical properties are difficult to predict.

The main design principles for the construction of a BDF nucleoside probe include effective mimicry of the parent nucleoside and highly sensitive fluorescent properties. We have previously developed a cytosine mimic, (^{DMA}**C**) that can fulfill both criteria in the context of folded nucleic acids. Here, we describe the ability of ^{DMA}**C** to serve as a BDF probe according to its distinct fluorescence quantum yield and anisotropy values when forming Watson-Crick versus mismatch pairs.

5.3. Results and Discussion

5.3.1. Oligonucleotide Sequences

To evaluate the ability of ^{DMA}**C** to detect single nucleotide mismatches in duplex DNA, the human telomeric DNA sequence modified with ^{DMA}**C** at two different positions was utilized. In these constructs, the fluorescent nucleoside ^{DMA}**C** was located in the center (**TeloX13**) or near the end (**TeloX2**) of the oligonucleotide.

5.3.2. Detecting Mismatches with ^{DMA}C

^{DMA}**C**-containing oligonucleotides were hybridized with complementary strands containing matched (G) or mismatched bases (C, T and A) opposed to ^{DMA}**C**. The resulting duplexes showed distinct photophysical properties according to the opposite and flanking bases (**Table 5.1**). The absorbance spectra of the mismatched duplexes were red-shifted (λ_{abs} = 372 – 402 nm) as compared to the matched sequences (λ_{abs} = 365 nm) and the emission spectra of all mismatched duplexes were centered at λ_{em} = 515 nm. In both constructs, the fluorescence quantum yields (ϕ) of ^{DMA}**C**-containing oligonucleotides were sensitive to the presence of matched or mismatched base pairing. With ^{DMA}**C** located in the center of the oligonucleotide (**TeloX13**), a two-fold decrease of fluorescence quantum yield was observed in the context of mismatched duplexes (ϕ = 0.05 – 0.07) as compared to the matched duplex (ϕ = 0.11, **Table 5.1** and **Figure 5.2**) The opposite trend was observed with ^{DMA}**C** located near the end of the oligonucleotide (**TeloX2**), with a two-fold increase of fluorescence quantum yield of the mismatched sequence (ϕ = 0.05 – 0.06, **Table 5.1** and **Figure 5.2**) as compared to the perfectly matched duplex (ϕ = 0.02). It is worth noting that the resulting fluorescence quantum yield in mismatched duplexes were very similar for both sequences.

In addition to the distinct properties in fluorescence emission, the fluorescence anisotropy (r) of the ^{DMA}**C**-containing oligonucleotides was highly characteristic to the presence of Watson-Crick base pairing. Indeed, mismatched pairing of ^{DMA}**C** with C, T and A should cause a significant increase of freedom of the fluorophore, as compared to base pairing with G. Consistent with this, the fluorescence anisotropy values of ^{DMA}**C** in perfectly matched (r = 0.14 – 0.15) was higher than mismatched base pairs (r = 0.09 – 0.10) for both DNA constructs evaluated. Together with the fluorescent anisotropy values measured previously

in unfolded DNA (r = 0.04 – 0.05), these results demonstrated that **^{DMA}C** is very sensitive to correct base paring in duplex DNA.

Sequence	Pairing	λ_{abs}[a]	λ_{em}[b]	ϕ[c]	ηt[d]	r[e]
TeloX13	**G**	365	505	0.11	0.04	0.14
X = ^{DMA}C	C	402	515	0.07	0.04	0.09
	A	372	515	0.07	0.06	0.09
	T	372	513	0.05	0.08	0.11
TeloX2	**G**	365	512	0.02	0.05	0.15
X = ^{DMA}C	C	392	515	0.06	0.06	0.09
	A	388	515	0.05	0.06	0.09
	T	390	515	0.05	0.06	0.11

Table 5.1. Photophysical data of ^{DMA}C in duplex DNA. [a] λ_{abs} are reported for the most red-shifted absorbance maxima, [b] λ_{em} in nm. [c] **^{DMA}C** (ϕ = 0.03 at pH = 7.35 with λ_{ex} = 370 nm) was used as the fluorescent standard for the relative quantum yields (ϕ) of **^{DMA}C** in DNA, [d] DNA to probe energy transfer efficiencies (ηt) were calculated at λ_{ex} = 260 nm and at the maximum λ_{em} (See the experimental section for calculation details), [e] fluorescence anisotropy (r) calculated at pH = 7.35 (λ_{ex} = 370 nm, λ_{em} = 535 nm). All measurements were performed with 4 μM of DNA in phosphate citric acid buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl)

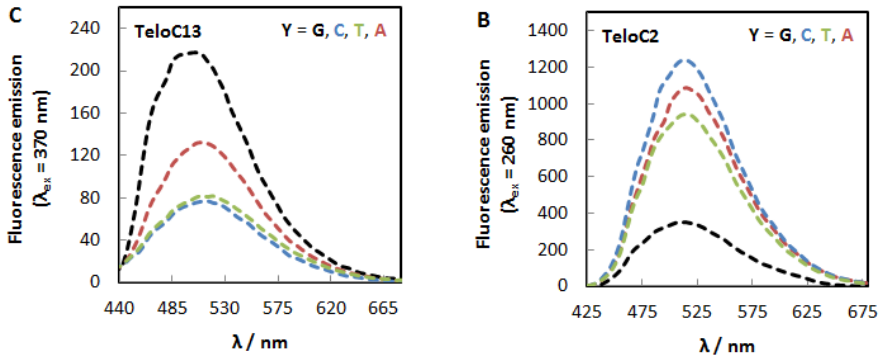


Figure 5.2. (A) Emission spectra (λ_{ex} = 370 nm) of **TeloX13** using selective excitation of the probe and (B) emission spectra (λ_{ex} = 260 nm) of **TeloX2** at pH = 7.35 (**X = ^{DMA}C**). DNAs were prepared as perfectly matched duplexes with C-G (black) or as mismatched duplexes with C-T (green), C-A (red) and C-C (blue).

5.4. Conclusions

The development of sensitive methods for the detection of SNPs is of high interest due to the relationships between genetic variations and complex diseases. With this ambition in mind, the fluorescent nucleoside analog ^{DMA}**C** was incorporated in the human telomeric DNA sequence and was evaluated in the presence of mismatched base pairs. The photophysical properties of ^{DMA}**C**-containing oligonucleotides, which include fluorescence quantum yields (ϕ = 0.02 – 0.11) and fluorescence anisotropy (r = 0.09 – 0.15) can be used to distinguish perfectly matched versus mismatched base pairs. ^{DMA}**C** can therefore serve as a base-discriminating fluorescent nucleoside analog for SNP typing.

Chapter 6 | A Fluorescent Thymidine Mimic for Probing Mercury(II)-Mediated Base Pairs

Chapter 6

A Fluorescent Thymidine Mimic for Probing Mercury(II)-Mediated Base Pairs

6.1. Summary

The interaction of Hg^{II} ions with thymidine-thymidine mismatches generates T–Hg^{II}–T base pairs that mimic Watson-Crick base pairs in duplex DNA. To specifically report the formation of T–Hg^{II}–T base pairs, a fluorescent nucleoside analog composed of a dimethylaniline unit, "^{DMA}**T**", was developed (**Figure 6.1**). Due to its effective mimicry of thymidine, ^{DMA}**T** has little or no impact on DNA stability but it exhibits well-resolved fluorescence properties indicative of DNA structure and base pairing. These include fluorescence anisotropy (r = 0.03 – 0.18) and quantum yield (ϕ = 0.03 – 0.20). These parameters were used to follow DNA strand displacement reactions using a single modified thymidine mimic. ^{DMA}**T** represents a promising internal probe that can innocently report the binding of Hg^{II} ions to thymidine in duplex DNA through dramatic changes in fluorescence quantum yields (ϕ = 0.01 – 0.13).

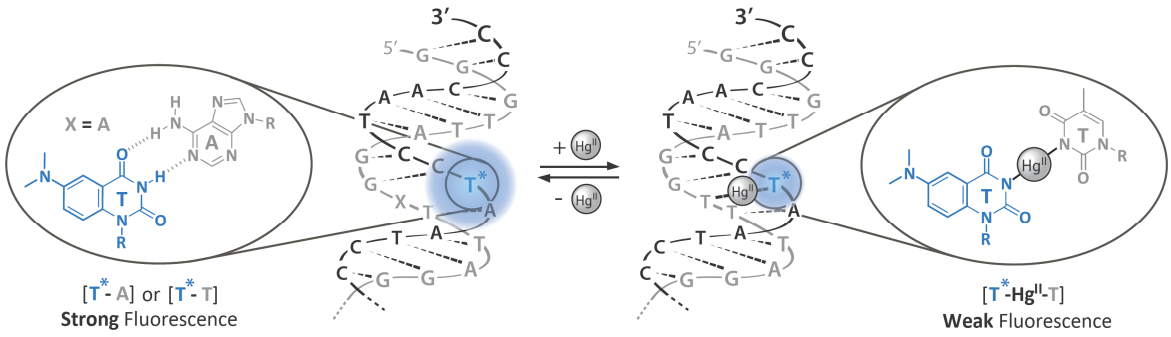


Figure 6.1. A fluorescent thymidine mimic (^{DMA}**T**) for Hg^{II}-mediated thymidine base pairs.

6.2. Introduction

Metal ions mediate a wide variety of biological processes in nucleic acids by maintaining the structural integrity of DNA and RNA.[298] The interaction of mercury with nucleic acid has been of particular interest due to the high cytotoxicity and mutagenicity of this heavy metal.[299-301] Hg^{II} ions have previously been shown to specifically bind to thymidine-thymidine mismatches in duplex DNA *via* the formation of T–Hg^{II}–T base pairs.[86] Among many metals known to interact with nucleic acids, only Hg^{II} was found to strongly bind thymidine mismatches,[86] at a molar ratio of 1:1 and with a reported binding constant of nearly 10^6 M^{-1}.[87] Direct evidence from NMR studies proved that the Hg^{II} ion coordinates to the *N3* positions of two thymidine residues following their deprotonation.[88] The formation of T–Hg^{II}–

T base pair resulted in stabilization of duplex DNA with thermal melting temperatures comparable to the corresponding Watson-Crick base pair.[89] The crystal and solution structure of the metallo-DNA complex were recently obtained. This has shown that the T–Hg^{II}–T base pairs are very well accommodated in duplex DNA and mimic of Watson-Crick base pairs.[90-91] Because of these features, DNA polymerase could misincorporate a thymidine triphosphate into the opposing site of a thymidine in the presence of Hg^{II} ions.[93]

A number of methods have been developed to detect the presence of T–Hg^{II}–T base pairs in duplex DNA. Biophysical techniques such as circular dichroism (CD)[86] provide global aspects of conformational changes, whereas NMR spectroscopy can provide highly detailed local parameters.[88] Fluorescence spectroscopy techniques provide the unique advantage of being compatible in living systems. End-labeled DNA have been used in FRET-based experiments to report conformation changes.[302-305] Although widely used, this method is limited to large changes in dye-to-dye distances, and can perturb the stability of DNA.[124-125] Since the native nucleosides in DNA are non-emissive,[110] synthetic fluorescent nucleosides[109] have been developed for monitoring base pairing interactions,[129,207] biochemical transformations,[130,192] and conformational changes.[75,250] However, only a small handful of fluorescent nucleoside analogs have been developed to probe metal binding in DNA.[306-307]

Herein, we describe a new fluorescent nucleoside analog (^{DMA}**T**) that reports the site-specific formation of T–Hg^{II}–T base pairs in duplex DNA. ^{DMA}**T** was very well accommodated in perfectly matched duplex DNA, where it exhibited higher emission quantum yield (ϕ = 0.03 – 0.20) as compared to the ^{DMA}**T** nucleoside in water (ϕ = 0.03). Upon addition of Hg^{II} ions to ^{DMA}**T**–T mismatched duplexes, the ^{DMA}**T**-containing oligonucleotides showed increases in thermal melting temperatures comparable to the wild type, T–T sequences. The large changes in fluorescence emission upon formation of ^{DMA}**T**–Hg^{II}–T base pair were used as a specific sensor for this highly toxic heavy metal.

6.3. A Fluorescent Thymidine Mimic

6.3.1. Probe Design

In order to report the formation of T–Hg^{II}–T base pairs in duplex DNA, we designed the nucleoside analog "*N*,*N*-dimethylaniline-2'-deoxythymidine", ^{DMA}**T** (**57**), containing a dimethylaniline unit fused to the pyrimidine ring (**Figure 6.2**). ^{DMA}**T** was designed to maintain global resemblance of the parent thymidine with respect to the overall structure, Watson-Crick hydrogen bonding face and ionization properties. Importantly, the *N*3-H acidity should be the same as thymidine to mimic the Hg^{II} binding site of T–T mismatches. The quinazoline scaffold has been previously shown to exhibit desirable photophysical properties and to be well accommodated into DNA structures.[114,130,207] To generate a push-pull fluorophore, an electron donating group was incorporated at the *C6* position of quinazoline. This position was selected over *C5* or *C7*, since *C6* is not directly conjugated with the *N*3-H position of thymidine, and should therefore have a minimal impact on its acidity. Molecular orbitals calculated using a DFT-optimized geometry of ^{DMA}**T** showed large charge transfer upon going from HOMO → LUMO (**Figure 6.3**), demonstrating that ^{DMA}**T** should exhibit "push-pull" fluorescence properties.

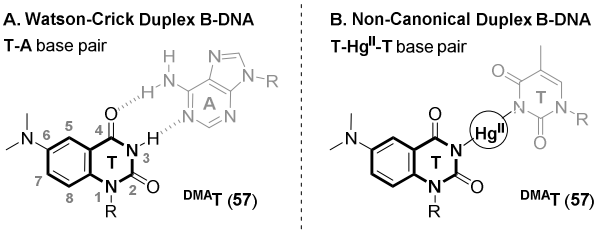


Figure 6.2. Structure of the designed thymidine mimic ^{DMA}**T** (**57**) engaged in (A) T-A Watson-Crick base pair and (B) in non-canonical T–Hg^{II}–T base pair within duplex B-DNA.

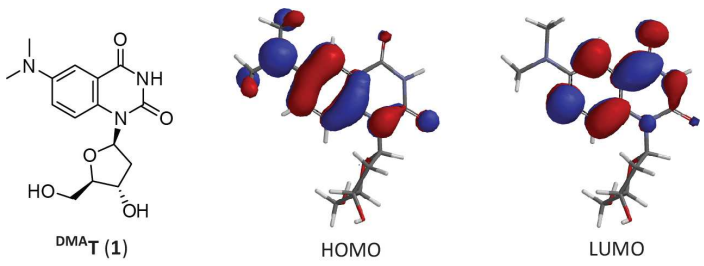


Figure 6.3. Structure of ^{DMA}**T** nucleoside (**57**), HOMO (E = -5.69 eV) and LUMO (E = -1.77 eV) molecular orbitals obtained from its DFT-optimized geometry using B3LYP/6-311++G**.

6.3.2. Synthesis and Solution Conformation of ^{DMA}T

The synthesis of the designed probe was performed using our previously reported route, including *N*-glycosylation of thioglycoside **14** and Buchwald-Hartwig coupling of intermediate **49** to afford compound **52** (**Figure 6.4**). Treatment with fluoride ions afforded the free nucleoside **57** (^{DMA}**T**) in 62% yield and high purity after washing the product several times with slightly acidic water to remove undesired tetrabutylammonium ions. To determine the impact of the dimethylaniline unit on the glycosidic bond orientation, a ^{1}H–^{1}H ROESY spectrum in DMSO-*d6* was collected (**Figure 6.5A**). Strong cross-peaks were found between H8 and H3 as well as H8 and H2β, which indicated the quinazoline core is located above the furanose ring. These correlations are consistent with an *anti*-conformation of the glycosidic bond. This conclusion is supported by the DFT-optimized structure (**Figure 6.5B**) which showed close proximities between H8 – H3' (2.58 Å) and H8 – H2'β (2.40 Å). Notably, another set of cross-peaks were found between H8 – H4' and H1' – H2'α, which confirmed the β-stereochemistry at the anomeric position. Together, these results demonstrate that ^{DMA}**T** possess the same conformation and stereochemistry as the parent thymidine.

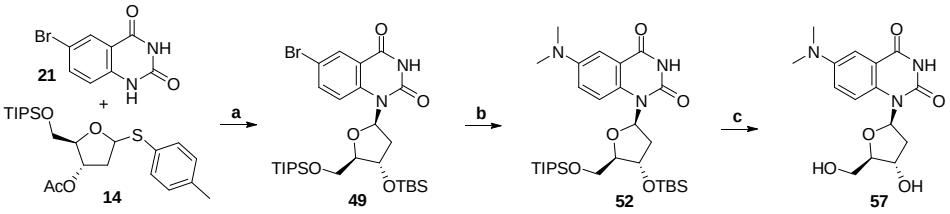


Figure 6.4. Synthesis of ^{DMA}**T** (**57**). Reaction conditions: a) (i) BSA, CH_2Cl_2, 23 °C, 2 h, then NIS, TMSOTf, 0 °C, 2 min, α/β = 1.0:1.8, 84% yield, (ii) K_2CO_3, MeOH then TBSCl, imidazole, DMF, 95% yield. b) Me_2NH in THF (2.0 M), $Pd_2(dba)_3$ (5 mol%), JohnPhos (20 mol%), KOtBu, dioxane, (0.15 M) 60 °C, 2 h, 82% yield. c) TBAF, THF, 23 °C, 2 h, 62% yield.

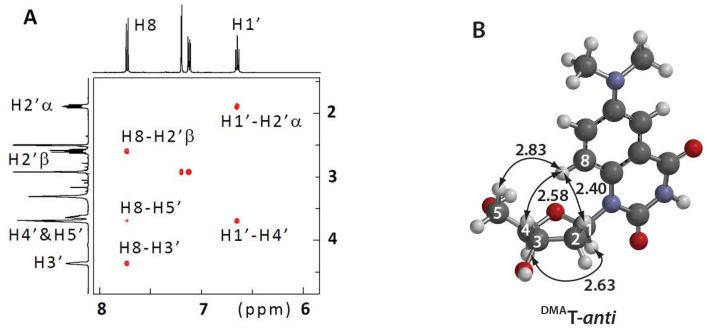


Figure 6.5. (A) Partial ^{1}H-^{1}H ROESY spectrum of nucleoside **57** (^{DMA}**T**) in DMSO-d_6 (mixing time of 250 ms). (B) Minimized structure using DFT calculation at the B3LYP/6-311G* level in Spartan 10.

6.3.3. Ionization Properties of ^{DMA}T

Next we analyzed the impact that the electron-rich dimethylaniline unit might have on the ionization properties of the pyrimidine ring. In order to measure the acidity of the *N*3-H, the absorbance and emission spectra of **^{DMA}T** were recorded at different pH values (pH = 7.55 to 11.5) in 0.1M phosphate citric acid buffer containing 0.1M NaCl (**Figure 6.6**). Interestingly, increased fluorescence intensities were observed upon deprotonation of the *N*3-H in water. This resulted in a 4-fold increase of quantum yield (ϕ = 0.12) as well as blue-shifted emission spectra (λ_{em} = 465 nm) at pH = 11.0 as compared to the neutral nucleobase (ϕ = 0.03 and λ_{em} = 522 nm) at pH = 7.35. The absorbance intensity at λ_{abs} = 345 nm and fluorescence intensity at λ_{em} = 470 nm (λ_{ex} = 360 nm) were plotted against the pH values. After fitting the data points, the pK_a of **^{DMA}T** was determined to be pK_a = 9.70 ± 0.05, which corresponds to the pK_a of thymidine. [31-32] These results demonstrate that the highly electron-rich dimethylaniline system has very little impact on the ionization of the *N*3 pyrimidine.

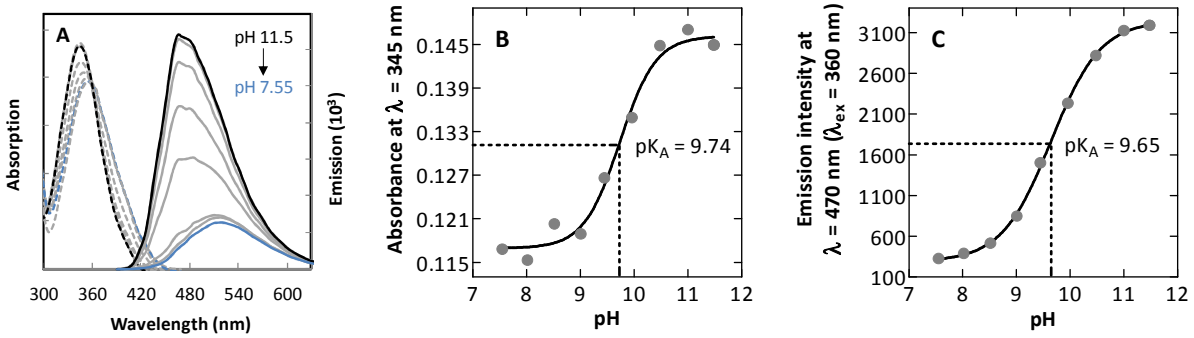


Figure 6.6. (A) Absorbance (---) and emission (—) spectra (λ_{ex} = 360 nm, isosbestic point) of 40 µM **^{DMA}T** (**57**) in an aqueous solution of 0.1M phosphate citric acid buffer containing 0.1M NaCl. (B) Plot between the absorbance (λ_{abs} = 345 nm) and the pH of the buffer. (C) Plot between the emission intensity (λ_{ex} = 360 nm and λ_{em} = 470 nm) and the pH of the buffer.

6.3.4. Environmental Sensitivity of ^{DMA}T

The photophysical properties of **^{DMA}T** are summarized in **Table 6.1**. The absorbance (λ_{abs} = 357 nm) and emission (λ_{em} = 522 nm) maxima of **^{DMA}T** are highly red-shifted as compared to native thymidine. A relative fluorescence quantum yield of ϕ = 0.03 was calculated using quinine hemisulfate (ϕ = 0.55 at λ_{ex} = 370 nm) as the fluorescent standard. In apolar solvents, such as dioxane, **^{DMA}T** showed highly increased quantum yields (ϕ = 0.60), as well as blue-shifted emission maxima (λ_{em} = 440 nm). In acetonitrile, a polar

organic solvent used as a mimic of DNA nucleobase environments, intermediate values of quantum yield ($\phi = 0.32$) and emission maxima ($\lambda_{em} = 458$ nm) were obtained. To further characterize the environmental sensitivity of ^{DMA}**T**, its absorption and emission properties were measured in water / dioxane mixtures (**Figure 6.7**). A linear correlation ($R^2 = 0.980$) with a large slope of 177 cm^{-1} / kcal mol^{-1} was obtained when plotting the Stokes shift against Reichardt's solvent polarity parameter (E_T^{30}).[219-220] In addition to the Stokes shift, the fluorescence intensities of ^{DMA}**T** were highly sensitive to the environment polarity. A linear correlation ($R^2 = 0.969$) with a slope of 0.023 mol kcal^{-1} was obtained when the fluorescence quantum yield was plotted against the E_T^{30} values. As previously observed with the related cytosine analog ^{DMA}**C**, these results demonstrate that ^{DMA}**T** is a push-pull fluorophore, where dimethylaniline acts as an electron donor and the pyrimidine acts as electron acceptor in the excited state.[229,260] This conclusion is further supported by DFT calculated molecular orbital plots of ^{DMA}**T**, which shows charge transfer from the electron-rich dimethylaniline to the electron-deficient pyrimidine ring upon going from HOMO → LUMO (**Figure 6.3**).

Solvent	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	$\varepsilon_{(260)}$[d]	$\varepsilon_{(\lambda abs)}$[d]	ϕ[e]
water	357	522	8.9	15.0	2.9	0.03
dioxane	372	440	4.2	14.1	4.0	0.60
acetonitrile	375	458	4.8	13.7	3.8	0.32
pH = 11.0	345	465	7.5	17.1	3.9	0.12

Table 6.1. Photophysical data of ^{DMA}**T** (**57**). [a] λ_{abs} are reported at the most red-shifted absorbance wavelength in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1}. [d] Extinction coefficients (ε) are reported in 10^3 M^{-1} cm^{-1} and are given at $\lambda_{abs} = 260$ nm and at the most red-shifted absorbance. [e] Quinine hemisulfate ($\phi = 0.55$ at $\lambda_{ex} = 370$ nm)[259] in 0.5 M H_2SO_4 was used as the fluorescent standard for calculating the relative quantum yields (ϕ). Associated errors are ±10% of the given values.

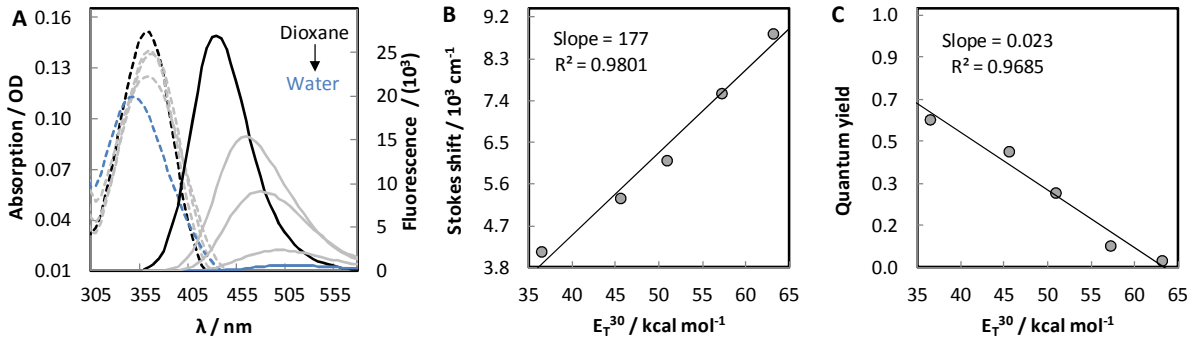


Figure 6.7. (A) Absorbance (---) and emission (—) spectra (λ_{ex} = 370 nm) of 40 μM ^{DMA}**T** (**1**) in dioxane (black), water (blue) and mixtures (grey). (B) Linear relationship between the Stokes shift and the E_T^{30}. (C) Linear relationship between the quantum yield and the E_T^{30}.

6.4. Probing DNA Structures with ^{DMA}T

6.4.1. Synthesis of ^{DMA}T-Containing DNA

To facilitate site-specific incorporation of ^{DMA}**T** into DNA, we prepared the corresponding phosphoramidite **59** suitable for standard automated DNA synthesis. For this aim, the primary alcohol of **57** was protected by 4,4'-dimethoxytrityl chloride (DMTCl), followed with phosphitylation of the secondary alcohol to afford phosphoramidite **59** in high yield (**Figure 6.8**).

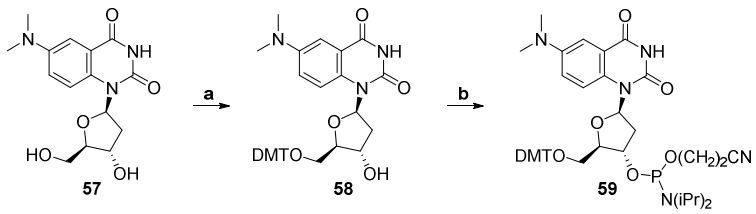


Figure 6.8. Synthesis of phosphoramidite **59**. Reaction conditions: d) DMTCl (1.2 equiv.), pyridine, 23 °C, 45 min, 79% yield. e) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (2.0 equiv.), DIPEA, CH_2Cl_2, 0 °C, 45 min, 86% yield.

We selected the cytosine rich human telomeric repeat sequence $(^{5'}CCCTAA^{3'})_n$ as a model for the incorporation of ^{DMA}**T** to allow comparisons with ^{DMA}**C** (See Chapters 5 and 6). ^{DMA}**T** was incorporated at four different positions within the same DNA sequence (**Figure 6.9**). The site-specific incorporation of the probe in oligonucleotides was confirmed using HR-MS of the HPLC-purified oligonucleotides.

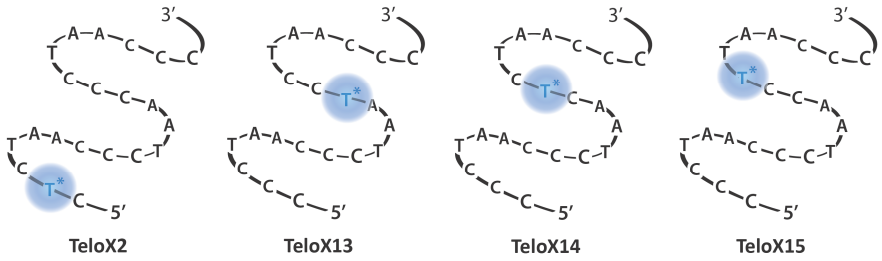


Figure 6.9. Oligonucleotides modified with ^{DMA}**T** at four different positions.

6.4.2. Impact of the Probe on Duplex Structure and Stability

Circular dichroism (CD) and thermal denaturation experiments were performed in order to assess the impact of ^{DMA}**T** on the structure and stability of duplex DNA (**Figure 6.10**, **Figure 6.11** and **Figure 6.12**). ^{DMA}**T**-containing oligonucleotides were therefore annealed to 1.1 equiv. of their perfect complementary strand by heating and slow cooling. For all constructs, a positive band at λ = 262 nm and a negative band at λ = 240 nm were observed in the resulting CD spectra, which is consistent with the formation of double-stranded helices (**Figure 6.10** and **Figure 6.11**).[95,261] Thermal denaturation melting temperatures (T_m) of ^{DMA}**T**-containing duplexes showed very little deviation (ΔT_m = -0.2 to -1.7 °C) as compared to wild-type sequences (**Figure 6.12** and **Table 6.2**). In contrast, when ^{DMA}**T** was replaced across from a mismatched nucleotide, such as C or A, large losses in thermal stabilities were observed (ΔT_m = -1.5 to -7.9 °C). These results demonstrated that ^{DMA}**T** has very little, if any, impact of the global structure or stability of duplex DNA, and that it exhibits the same base pairing preferences as thymidine.

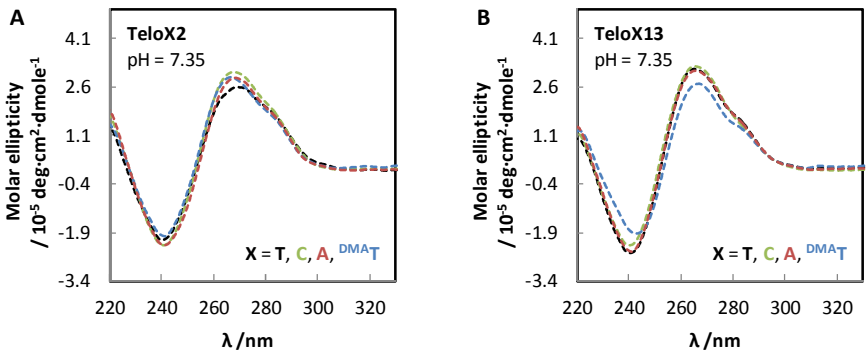


Figure 6.10. CD spectra at 25 °C of duplex DNA containing (A) **TeloX2**, (B) **X13**, (C) **X14** or (D) **X15** with **X** = **T**, **C**, **A** and ^{DMA}**T**.

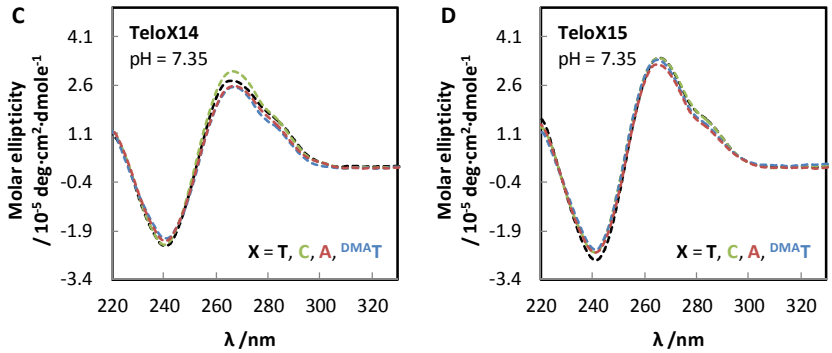


Figure 6.11. CD spectra at 25 °C of duplex DNA containing (A) **TeloX2**, (B) **X13**, (C) **X14** or (D) **X15** with **X** = **T**, **C**, **A** and ^{DMA}**T**. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) at pH = 7.35. Double-stranded DNA samples were prepared using 1.1 equivalents of the complementary strand containing an "**A**" residue opposite to the modification site "**X**".

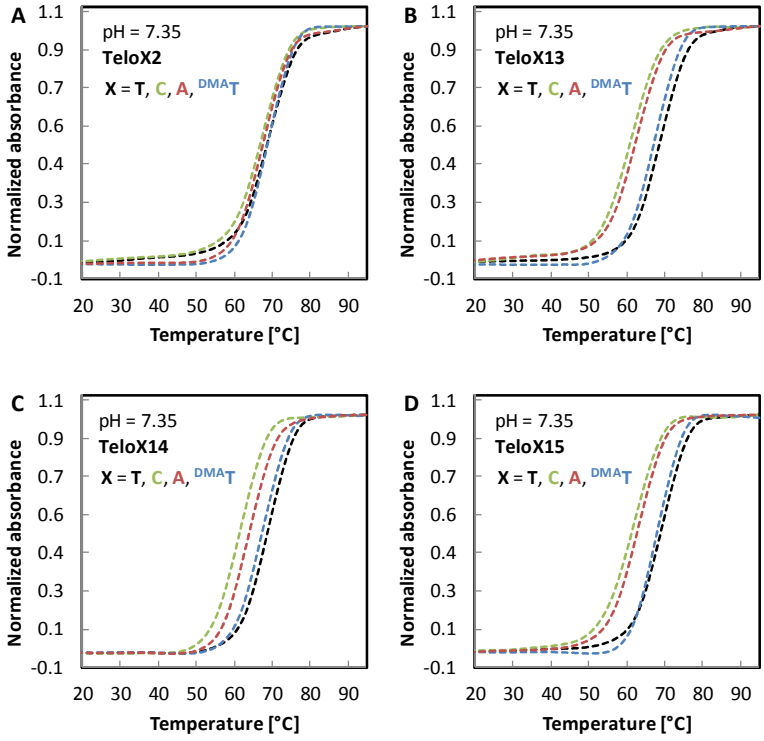


Figure 6.12. Thermal denaturation curves of duplex DNA containing (A) **TeloX2**, (B) **X13**, (C) **X14** or (D) **X15** with **X** = **T**, **C**, **A** and ^{DMA}**T**. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) at pH = 7.35. Double-stranded DNA samples were prepared using 1.1 equivalents of the complementary strand containing an "**A**" residue opposite to the modification site "**X**".

	T_m of duplex DNA			
X	**T (wt)**	**^{DMA}T**	**C**	**A**
TeloX2	69.0	68.8 (-0.2)	67.3 (-1.7)	67.5 (-1.5)
TeloX13	68.8	67.1 (-1.7)	60.9 (-7.9)	62.4 (-6.4)
TeloX14	68.7	67.3 (-1.4)	61.3 (-7.4)	63.9 (-4.8)
TeloX15	68.9	68.0 (-0.9)	61.2 (-7.7)	62.7 (-6.2)

Table 6.2. Thermal denaturation melting temperature (T_m) of duplex DNA containing **TeloX2**, **X13**, **X14** or **X15** with **X** = **T**, **C**, **A** and **^{DMA}T** and deviation from wild-type (+/-ΔT_m). All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) at pH = 7.35. Double-stranded DNA samples were prepared using 1.1 equivalents of the complementary strand containing an "**A**" residue opposite to the modification site "**X**". Reproducibility is within 0.5 °C.

6.4.3. Assignment of Static Structure with ^{DMA}T

The fluorescence properties of **^{DMA}T** are sensitive to the folded structure of the DNA (**Table 6.3** and **Figure 6.13** and **Appendix A14**). Notably, **^{DMA}T** exhibited increased fluorescence quantum yield (ϕ = 0.03 – 0.20) in DNA as compared to the free nucleoside in water (ϕ = 0.03). An exceptionally high ~7-fold increase was observed when **^{DMA}T** was flanked by adenosine (**TeloX13**). Unlike **TeloX2**, where **^{DMA}T** is located near the end of the oligonucleotide and suffers from end-breathing effects, all other duplexes showed a 2-fold higher fluorescence quantum yield as compared to unfolded DNA (**Figure 6.13**). The absorbance and emission (λ_{abs} = 355 nm and λ_{em} = 486 – 505 nm) of **^{DMA}T** in duplex DNA are blue-shifted as compared to unfolded DNA (λ_{abs} = 365 nm and λ_{em} = 515 – 516 nm). Similar trends were observed for the **^{DMA}T** nucleoside in organic solvents (**Table 6.1**), suggesting that **^{DMA}T** was protected from bulk water in duplex DNA. A large change in fluorescence anisotropy was observed in the unfolded (r = 0.03 – 0.06) versus duplex DNA (r = 0.09 – 0.18). The excitation spectra of the fluorescent probe in oligonucleotides exhibited broad maxima at λ_{ex} = 260 nm and λ_{ex} = 365 nm, which revealed energy transfer from the unmodified bases (λ_{em} = 327 nm)[264] to **^{DMA}T** (**Figure 6.13**).[265-267] Quantification of energy transfer efficiencies revealed a ~2-fold decrease in efficiency in unfolded versus duplex DNA. Taken together, these results demonstrated that **^{DMA}T** can serve as a sensitive reporter of DNA structure.

Sequence	Folding	λ_{abs} [a]	λ_{em} [b]	r [c]	ϕ [d]	η_t [e]
TeloX2	Unfolded	365	515	0.03	0.05	0.06
C-X-C	Duplex	355	505	0.18	0.03	0.04
TeloX13	Unfolded	365	515	0.05	0.09	0.08
A-X-C	Duplex	355	504	0.09	0.20	0.04
TeloX14	Unfolded	365	516	0.06	0.06	0.04
C-X-C	Duplex	355	492	0.14	0.13	0.02
TeloX15	Unfolded	365	515	0.05	0.05	0.07
C-X-T	Duplex	355	486	0.15	0.11	0.05

Table 6.3. Photophysical data of emissive nucleoside **1** (^{DMA}**T**) in DNA (**X** = ^{DMA}**T**). [a] λ_{abs} are reported at the most red-shifted absorbance maxima in nm, [b] λ_{em} in nm, [c] Fluorescence anisotropy (*r*) were calculated at pH = 7.35 (λ_{ex} = 375, nm λ_{em} = 500 nm). [d] ^{DMA}**T** (ϕ = 0.03 at pH = 7.35 with λ_{ex} = 355 nm) was used as the fluorescent standard for calculating the relative quantum yields (ϕ) of ^{DMA}**T** in DNA. [e] DNA-to-probe energy transfer efficiencies (ηt) were calculated at λ_{ex} = 260 nm. All samples contained 4 μM of DNA in phosphate citric acid buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl at pH = 7.35).

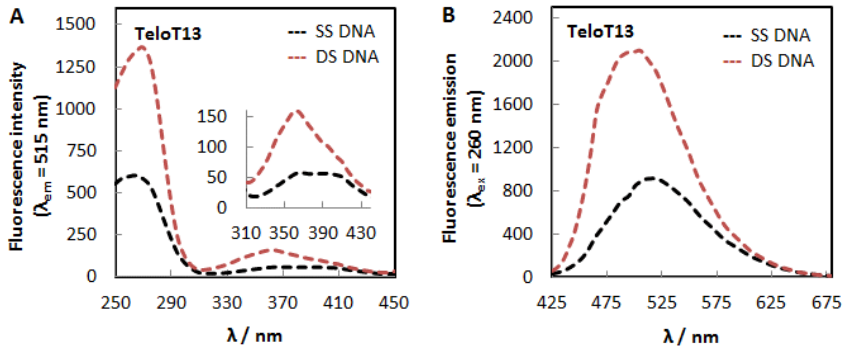


Figure 6.13. (A) Excitation (λ_{em} = 515 nm) and (B) fluorescence spectra (λ_{ex} = 260 nm) of **TeloX13**, prepared as unfolded (black, SS) or double-stranded (red, DS) DNA. All samples contained 4 μM of DNA in phosphate citric acid buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl at pH = 7.35).

6.4.4. Probing Real-Time DNA Dynamics

To evaluate the ability of ^{DMA}**T** to serve as a real-time probe of DNA conformational changes, we conducted DNA strand displacement assays.[271] ^{DMA}**T**-containing DNA were pre-hybridized with a complementary strand containing a variable-length "toehold" (TH) to give "**GC**" duplexes with a short, 5'-single-stranded overhang. Thermodynamically favorable displacement of the labeled strand is accomplished by adding an excess of an unlabeled invading strand "**I**" to form a new duplex "**GI**" that contains no single-stranded overhang (**Figure 6.14**).

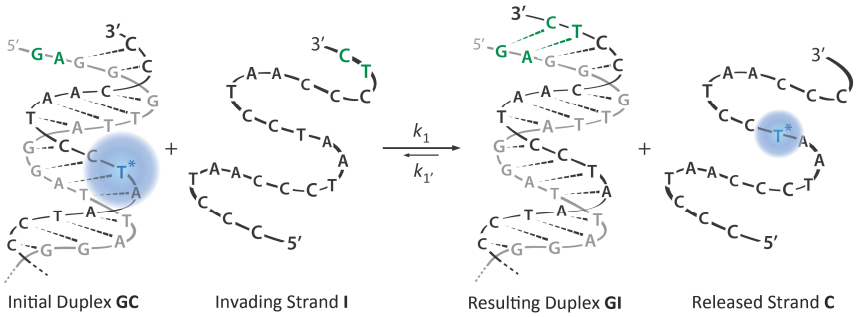


Figure 6.14. Representation of strand displacement reaction at pH = 7.35 with **TeloX13** (**X** = ^{DMA}**T**) and a toehold of 2 nucleotides (green).

Strand displacement was monitored using the large difference in fluorescence quantum yield of ^{DMA}**T** in duplex versus unfolded structures (**Figure 6.13**). Rate constants (k) were calculated using pseudo-first approximations as previously described,[268-270] with increasing concentrations of invading strand **I** ranging from four- to a ten-fold excesses (**Figure 6.15**). After fitting the data to single-exponential decays, the observed rates (k') were plotted against the initial invading strand concentration ($\mathbf{I_0}$) to give rate constant (k) as the slope of each plot. As expected, rates constants increased with increasing toehold lengths, $k_{TH1} = 0.81 \times 10^1$ ($R^2 = 0.978$) < $k_{TH2} = 2.4 \times 10^1$ M^{-1} s^{-1} ($R^2 = 0.990$). These values are very similar to the rates obtained from DNA strand displacements experiment conducted using ^{DMA}**C** (See Chapter 4) or FRET-labels.[271] When ^{DMA}**T** is opposed to a thymidine instead of an adenosine in the initial duplex **GC**, a 2-fold faster rate was observed, $k_{TH2} = 5.3 \times 10^1$ ($R^2 = 0.968$). These results validate ^{DMA}**T** as a probe for monitoring real-time DNA dynamics in perfectly matched and mismatched duplexes.

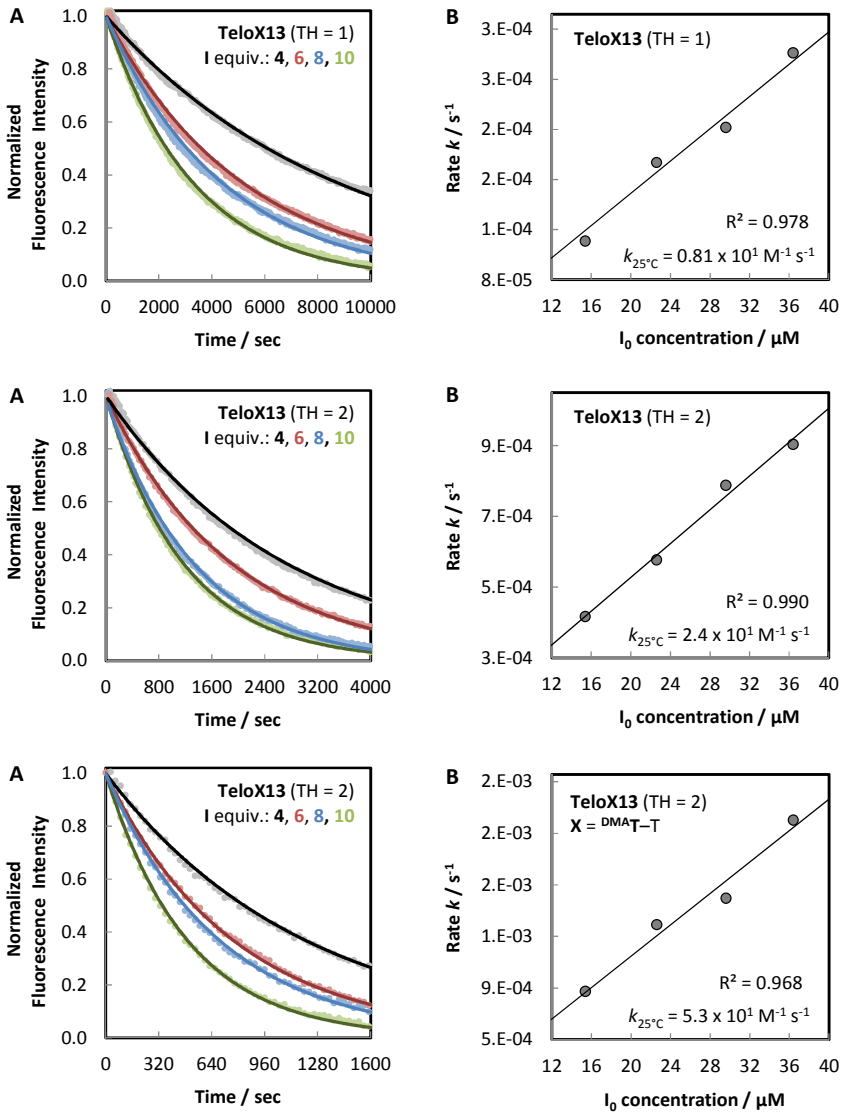


Figure 6.15. (A) Real-time strand displacement according to fluorescence intensity of **TeloX13** (**X** = ^{DMA}**T**, λ_{ex} = 370 nm and λ_{em} = 500 nm) with toehold (TH) = 1 and 2. The lines represent the mono-exponential fit of the data to eq. 4. (B) Rates (s^{-1}) plotted against $\mathbf{I_0}$ concentration (µM) gave the rate constant (M^{-1} s^{-1}) after linear fitting. Rate constants k_1 were calculated by fitting of fluorescence intensity data (λ_{ex} = 370 nm, λ_{em} = 500 nm) to a single exponential decay eq. 4. Measurements were performed using 4 µM of DNA in phosphate citric acid (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) buffer at pH = 7.35. The oligonucleotide sequences of each strand are provided in the Appendix.

6.5. Probing T–Hg^{II}–T in DNA

6.5.1. Site-Selective Formation of T–Hg^{II}–T Base Pair

After having validated ^{DMA}**T** as a sensitive fluorescent mimic of thymidine, we then turned our attention to the formation of the T–Hg^{II}–T base pair within duplex DNA. The wild-type sequence **TeloT13** was annealed to the complementary sequence containing a T mismatch in 0.1M phosphate citric acid buffer containing 0.1M $NaNO_3$ (T_m = 61.5 °C and ΔT_m = -7.3 °C as compared to the perfect complement, **Table 6.4**). We then measured the melting temperatures of this duplex upon addition of Hg^{II} ions. After the addition of 1.1 equiv. of $Hg(ClO_4)_2$, a large increase in thermal stability was observed for the T–T mismatched oligonucleotide (T_m = 67.5 °C and ΔT_m = +6.0 °C, **Figure 6.16**). This melting temperature indicates a similar stability as the T–A base pair in the perfect duplex (T_m = 68.8 °C), confirming the resemblance of the T–Hg^{II}–T to Watson-Crick base pairs. Further addition of Hg^{II} ions (until 2.0 equivalents) resulted in very little further increases of the T_m values, demonstrating the site-selective incorporation of a single Hg^{II} ion into the T–T site. These results were similar for all four DNA sequences studied (ΔT_m = +1.5 to +6.0 °C, **Table 6.4**). Moreover, CD spectra of the DNA gave noticeable changes in molar ellipticity intensities in the presence of Hg^{II} ions (**Figure 6.18**). In stark contrast, the addition of Hg^{II} ions in the C–T mismatched pairs showed only negligible changes in melting temperatures (ΔT_m = -0.5 to +0.7 °C) and molar ellipticity changes. Taken together, these results demonstrated the site-selective incorporation of Hg^{II} ions to form a T–Hg^{II}–T base pair in duplex DNA.

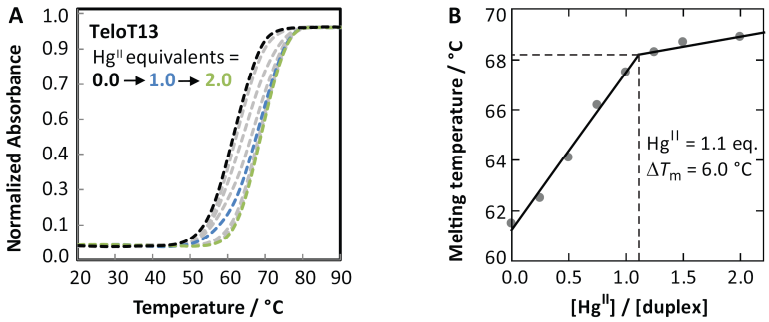


Figure 6.16. (A) Thermal denaturation curves of duplex **TeloT13** in the absence (black) and in the presence of $Hg(ClO_4)_2$ (1.0 equiv. (blue), 2.0 equiv. (green) and intermediate (grey)). (B) Plot of the melting temperature versus equivalents of Hg^{II} ions. The complementary strand contained a "**T**" mismatched opposite to the "**T13**" base. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35.

We then investigated the relative impact of the thymidine mimic ^{DMA}**T** on the stabilities of T–T mismatched and T–Hg^{II}–T base pairs. In the absence of Hg^{II} ions, small thermal stabilizations (ΔT_m = +0.8 to +1.2 °C, **Table 6.4** and **Figure 6.17**) were observed in ^{DMA}**T**–T versus T–T mismatches, probably due to the enhanced stacking interactions of the dimethylaniline unit with flanking nucleobases.[115] Upon addition of 1.0 equiv. of Hg^{II} ions, increases in thermal stabilization (ΔT_m = +0.8 to +4.8 °C, **Table 6.4**) were observed, consistent with the formation of ^{DMA}**T**–Hg^{II}–T base pairs. Very similar T_m values were obtained as compared to the wild-type sequences (ΔT_m = -0.4 to +1.2 °C), which showed that ^{DMA}**T** has very little, if any impact on the global stability of duplex DNA containing ^{DMA}**T**–Hg^{II}–T base pairs. In addition, CD spectra of the ^{DMA}**T** containing oligonucleotides exhibited very similar behavior as compared to the wild-type sequences upon addition of 1.0 equiv. of Hg^{II} ions (**Figure 6.18**). Taken together, these results demonstrate the excellent mimicry of ^{DMA}**T** for thymidine residues in duplexes containing T–A and T–Hg^{II}–T base pairs.

T_m of duplex DNA in the absence or presence of Hg^{II}						
Base	**T**		^{DMA}**T**		**C**	
Sequence	-	+ Hg^{II}	-	+ Hg^{II}	-	+ Hg^{II}
TeloX2	67.0	68.5 (+1.5)	68.0	68.8 (+0.8)	66.8	66.3 (-0.5)
TeloX13	61.5	67.5 (+6.0)	62.3	67.1 (+4.8)	60.4	61.1 (+0.7)
TeloX14	61.5	65.1 (+3.6)	62.5	65.7 (+3.2)	60.4	60.3 (-0.1)
TeloX15	62.8	65.7 (+2.9)	64.0	66.9 (+2.9)	60.6	60.9 (+0.3)

Table 6.4. Thermal denaturation melting temperature (T_m) of duplex of **TeloX2**, **X13**, **X14** and **X15** with **X** = **T**, **C** and ^{DMA}**T** in the absence and in the presence (+/-ΔT_m) of Hg^{II} ions (1.0 equiv.). The complementary strand contained a T mismatched opposite to the **X** base. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35.

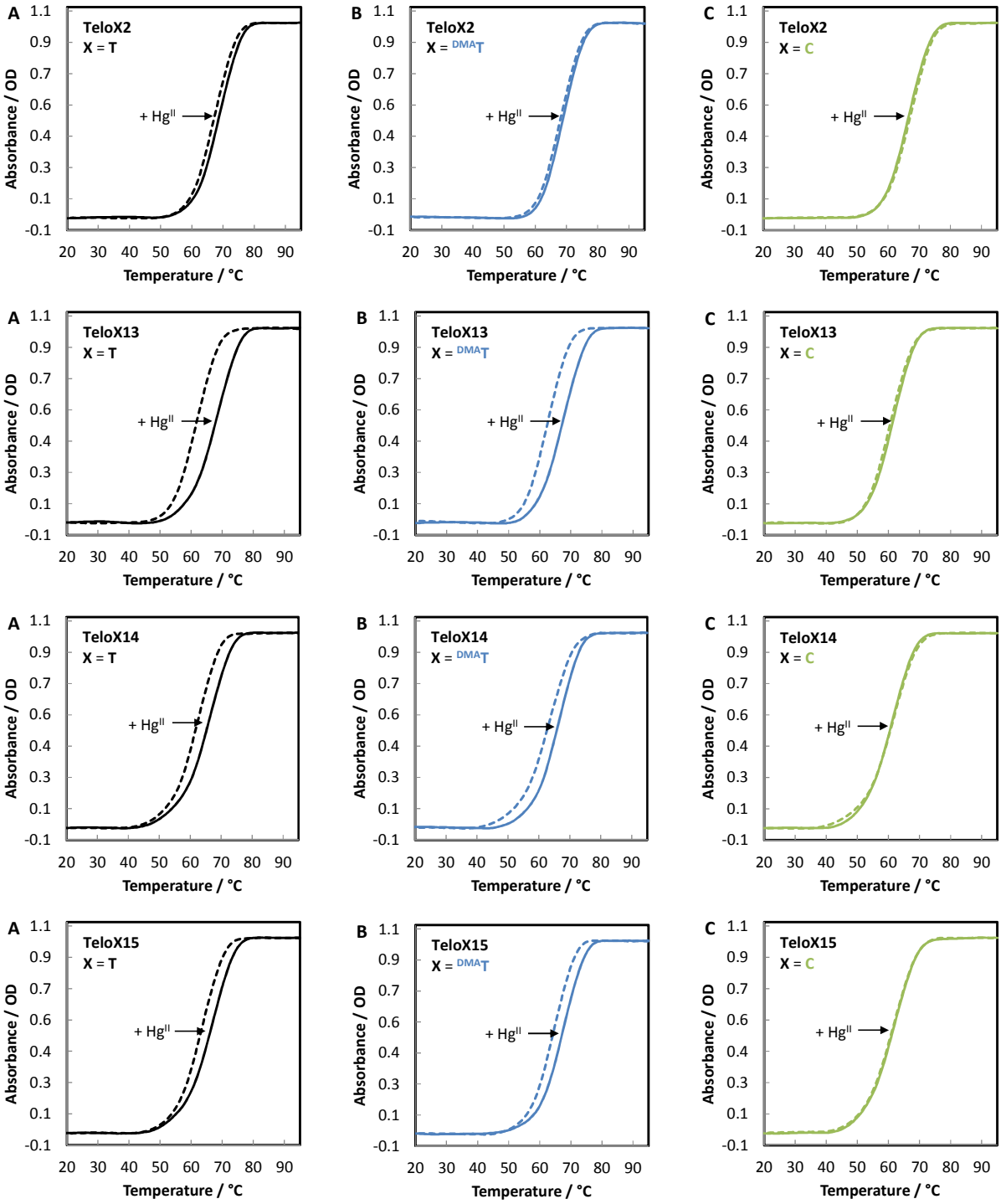


Figure 6.17. Thermal denaturation curves of duplex DNA containing **TeloX2**, **TeloX13**, **TeloX14** or **TeloX15** with **X** = (A) **T**, (B) ^{DMA}**T** and (C) **C** in the absence (---) and in the presence (—) of $Hg(ClO_4)_2$ (1.0 equiv.). The complementary strand a "**T**" mismatched opposite to the "**X**" base. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35.

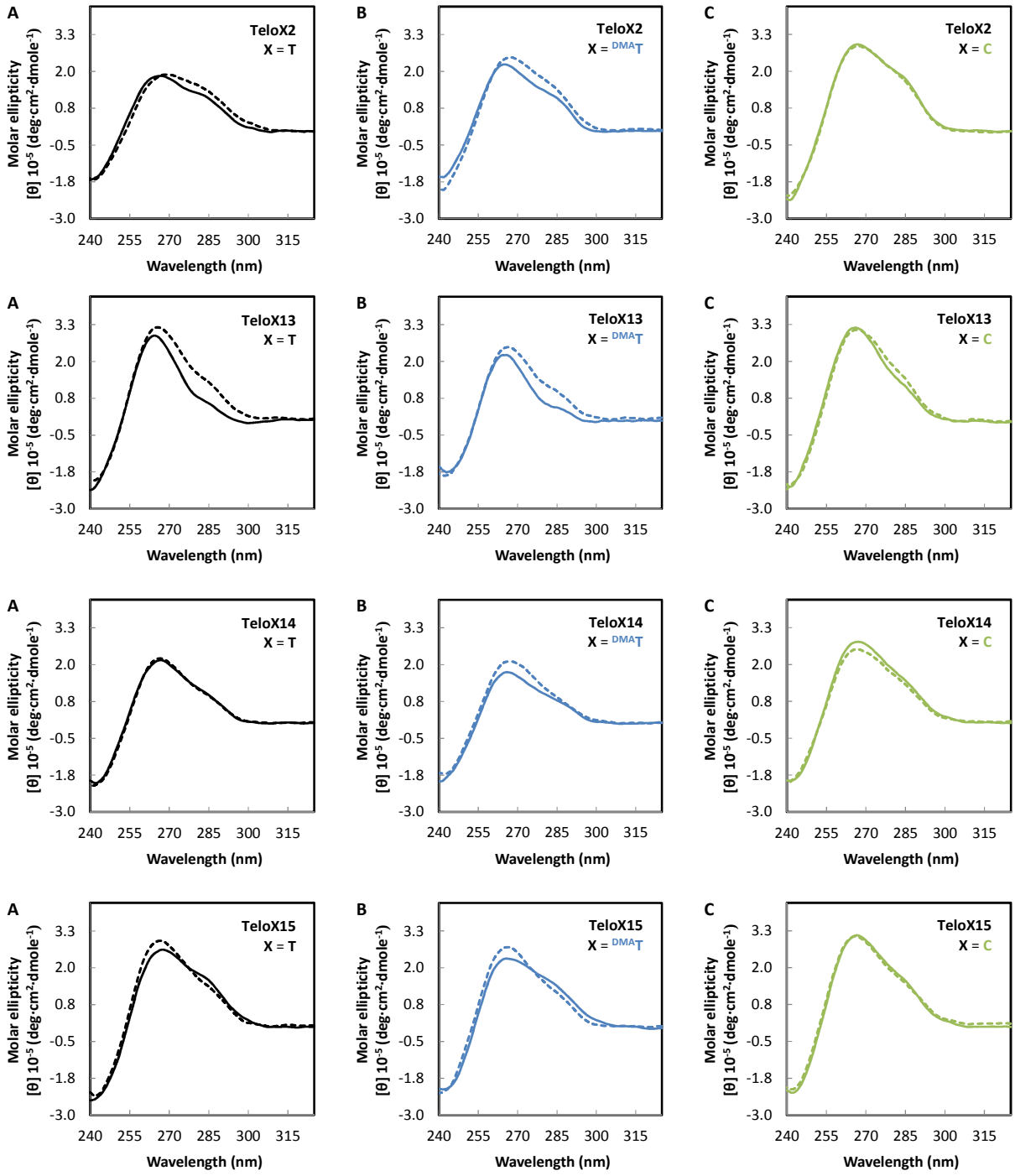


Figure 6.18. CD spectra of duplex DNA containing **TeloX2**, **TeloX13**, **TeloX14** or **TeloX15** at 5 °C with **X** = (A) **T**, (B) ^{DMA}**T** and (C) **C** in the absence (---) and in the presence (—) of $Hg(ClO_4)_2$ (1.0 equiv.). The complementary strand contained a "**T**" mismatched opposite to the "**X**" base. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35.

6.5.2. Probing T–Hg^{II}–T Base Pair Formation with Fluorescence

The fluorescence properties of ^{DMA}**T** were used to follow the incorporation of Hg^{II} ions in T–T pairs. ^{DMA}**T**-containing oligonucleotides were annealed to their complementary sequences to form ^{DMA}**T**–T mismatches. The resulting duplexes exhibited strong blue-shifted emission centered at λ_{em} = 500 – 510 nm and enhanced fluorescence quantum yield (ϕ = 0.05 – 0.13) as compared to the free nucleoside (**57**) in water (λ_{em} = 522 nm and ϕ = 0.03) (**Table 6.5** and **Figure 6.21**). Upon addition of 1.35 equivalents of Hg^{II} ions, a 5-fold decrease in fluorescence intensity of the duplexes was observed (**Figure 6.19**). The fluorescence quantum yield measured at 1.0 equiv. of Hg^{II} ions was ϕ = 0.05. Similar results were obtained when ^{DMA}**T** was located at all four different positions within the DNA sequence (**Figure 6.20** and **Figure 6.21**). These results demonstrate that ^{DMA}**T** can be used to probe the site-selective insertion of Hg^{II} ions at various T–T mismatch according to its large changes in fluorescence emission.

Fluorescence anisotropy (r) is a powerful tool to measure the dynamic motions of fluorophores.[108] Upon coordination of Hg^{II} ions to ^{DMA}**T**, very little difference in fluorescence anisotropy was observed as compared to the T–T mismatch (Δr = -0.01 – +0.02), which suggested that the dynamic motion of T–Hg^{II}–T base pair is comparable to that of T–T mismatch, due to the lack of hydrogen bonds. This result is consistent with the cationic character of Hg^{II} ions in T–Hg^{II}–T base pairs, determined by Raman spectroscopy and DFT calculations,[308] which showed that the *N3*-Hg^{II} bond is less covalent and more ionic that the *N3*-H bond.

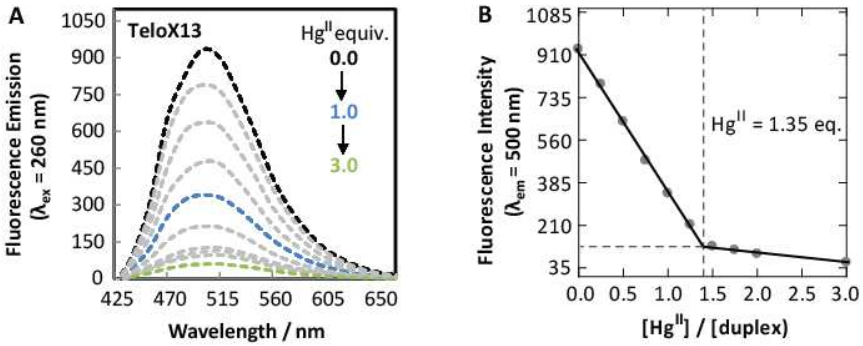


Figure 6.19. (A) Fluorescence spectra (λ_{ex} = 260 nm) of duplex **TeloX13** in the absence (black) and in the presence of variable equivalents of $Hg(ClO_4)_2$. (B) Plot of fluorescence intensity (λ_{em} = 500 nm) versus equivalents of Hg^{II} ions. The complementary strand contained a "**T**" mismatched opposite to the "**X13**" base. All samples contained 2 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35. All samples were heated to 90 °C for 5 min and cooled to rt overnight before reading.

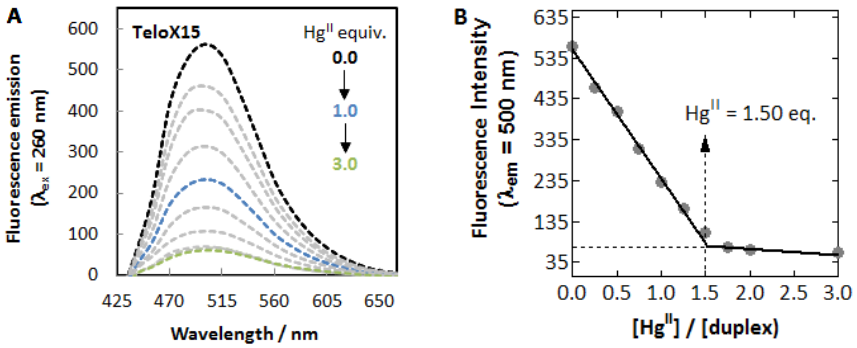


Figure 6.20. (A) Fluorescence spectra (λ_{ex} = 260 nm) of duplex **TeloX15** in the absence (black) and in the presence of variable equivalents of $Hg(ClO_4)_2$. (B) Plot of fluorescence intensity (λ_{em} = 500 nm) versus equivalents of Hg^{II} ions. The complementary strand contained a "**T**" mismatched opposite to the "^{DMA}**T15**" base. All samples contained 2 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35. All samples were heated to 90 °C for 5 min and cooled to rt overnight before reading.

Sequence	Base Pairing	λ_{abs}[a]	λ_{em}[b]	r[c]	ϕ[d]	η_t[e]
TeloX2	^{DMA}T-/-T	375	510	0.10	0.05	0.07
C-T-C	^{DMA}T-Hg^{II}-T	360	512	0.10	0.01	0.10
TeloX13	^{DMA}T-/-T	365	505	0.09	0.13	0.08
A-T-C	^{DMA}T-Hg^{II}-T	360	507	0.09	0.05	0.07
TeloX14	^{DMA}T-/-T	385	501	0.13	0.07	0.07
C-T-C	^{DMA}T-Hg^{II}-T	360	504	0.15	0.03	0.06
TeloX15	^{DMA}T-/-T	365	500	0.12	0.08	0.08
C-T-T	^{DMA}T-Hg^{II}-T	360	505	0.11	0.03	0.09

Table 6.5. Photophysical data of ^{DMA}**T** in duplex DNA. [a] λ_{abs} are reported at the most red-shifted absorbance maxima in nm, [b] λ_{em} in nm, [c] Fluorescence anisotropy (r) were calculated at pH = 7.35 (λ_{ex} = 375 nm. λ_{em} = 500 nm). [d] ^{DMA}**T** (ϕ = 0.03 at pH = 7.35 with λ_{ex} = 370 nm) was used as the fluorescent standard for calculating the relative quantum yields (ϕ) of ^{DMA}**T** in DNA. [e] DNA-to-probe energy transfer efficiencies (ηt) were calculated at λ_{ex} = 260 nm. T–Hg^{II}–T base pairs were generated by the addition of 1.0 equiv. of $Hg(ClO_4)_2$ followed by heating to 90 °C for 5 min and cooled to rt overnight before reading.

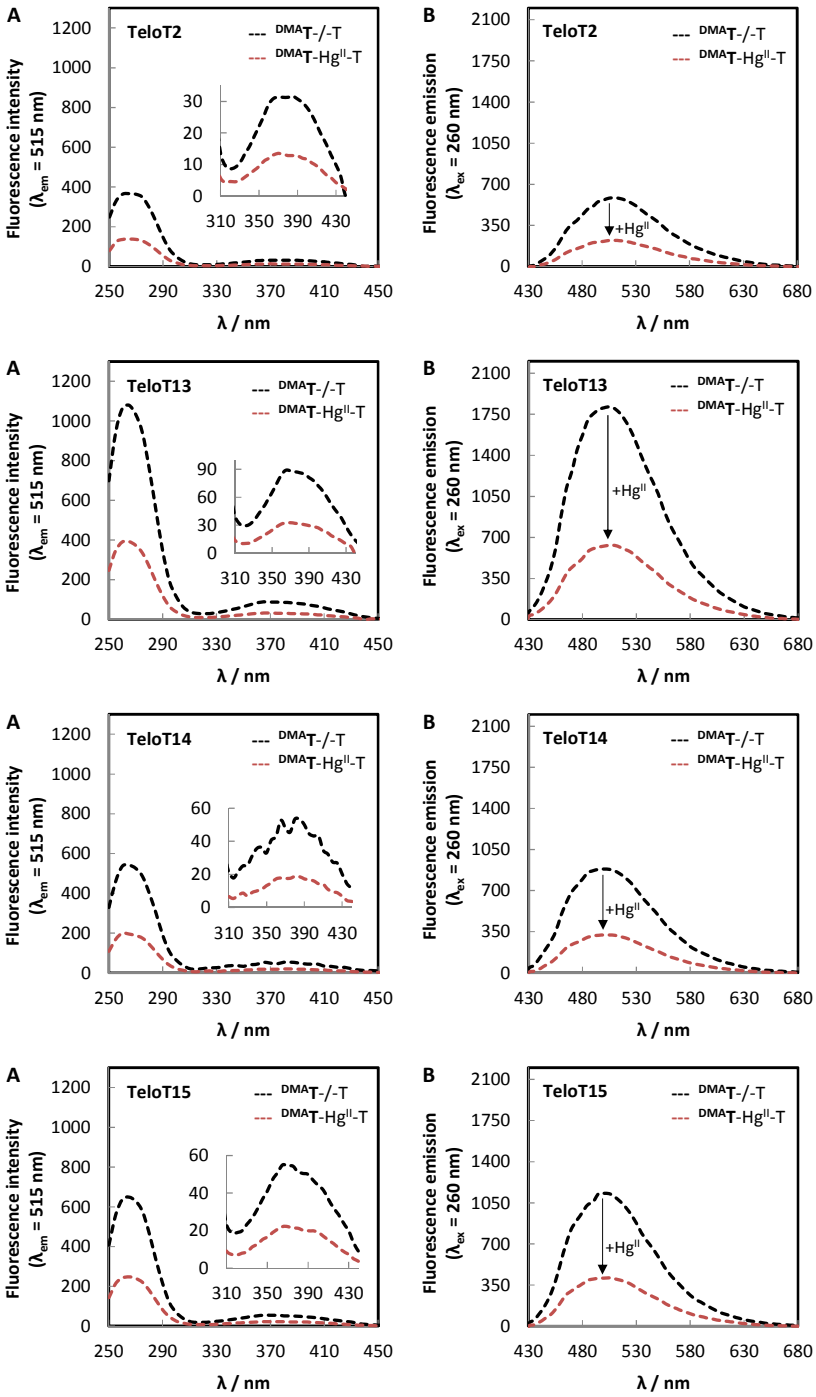


Figure 6.21. (A) Excitation (λ_{em} = 515 nm) and (B) fluorescence spectra (λ_{ex} = 260 nm) of **TeloT2**, **TeloT13**, **TeloT14** and **TeloT15** as double-stranded DNA with a T–T mismatch (black) and a T-Hg^{II}-T match (red). All samples contain 4 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35. The complementary strand contained a "**T**" mismatched opposite to the "^{DMA}**T**" base. T–Hg^{II}–T base pairs were generated by the addition of 1.0 equiv. of $Hg(ClO_4)_2$ followed by heating to 90 °C for 5 min and cooled to rt overnight before reading.

6.5.3. Metal Selectivity

We then evaluated the selectivity of ^{DMA}**T** for Hg^{II} ions in duplex DNA containing one ^{DMA}**T**–T mismatch pair. ^{DMA}**T**-containing duplexes (**TeloX13** and **TeloX15**) were therefore prepared by heating and slow cooling the oligonucleotides in the presence of metal cations known to coordinate to nucleic acids, such as Zn^{II}, Cu^{II}, Mg^{II}, Fe^{II}, Ca^{II}, Ag^{I}, Cd^{II}, Pd^{II} and Ni^{II}. Unlike Hg^{II}, these metal ions had very little impact on the fluorescence quantum yield of ^{DMA}**T** for both DNA constructs tested (**Figure 6.22**). This results suggested that these metal cations are not be incorporated into ^{DMA}**T**–T mismatched pairs or that they had no impact on the fluorescence of ^{DMA}**T**-containing duplexes. This conclusion is consistent with thermal denaturation experiments that showed only negligible effects of these metal cations on melting temperatures of T–T mismatches.[86] This results demonstrate that ^{DMA}**T** can serve as a specific fluorescent sensor for Hg^{II} ions.

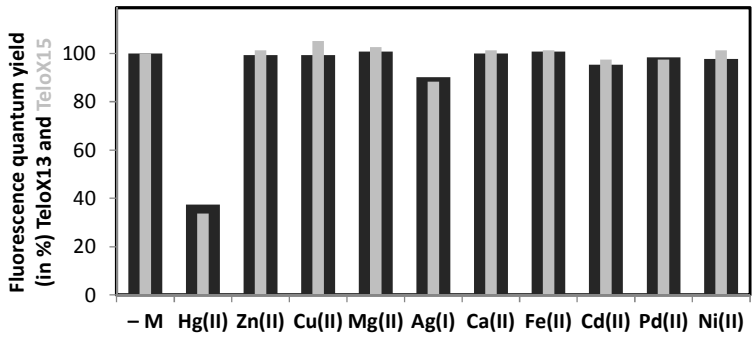


Figure 6.22. Fluorescence quantum yield of ^{DMA}**T** in T–T mismatched duplex DNA with **TeloX13** (black) and **TeloX15** (grey) upon addition of metal ions (1.0 equiv.). $Hg(ClO_4)_2$, $ZnCl_2$, $CuCl_2$, $MgCl_2$, $FeCl_2$, $CaCl_2$, $CdCl_2$, $PdCl_2$, $NiCl_2$, $AgNO_3$ were used in these experiments. All samples contained 4 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl or $NaNO_3$) at pH = 7.35.

6.5.4. Base Pairing Effects

We then turned our attention to the effect of Hg^{II} ions on duplex DNA containing ^{DMA}**T** paired to A, G and C. Little impact on the fluorescence quantum yield of ^{DMA}**T** was observed when Hg^{II} ions were added to duplex containing purine nucleobases A and G (**Figure 6.23**). When ^{DMA}**T** was paired to the pyrimidine nucleobase C, a more pronounced impact on the fluorescence of ^{DMA}**T** was observed, suggesting some coordination of Hg^{II} ions between cytosine and ^{DMA}**T**. In contrast to this, thermal denaturation melting experiments could not identify any effect of Hg^{II} ions on C–T stability.[89]

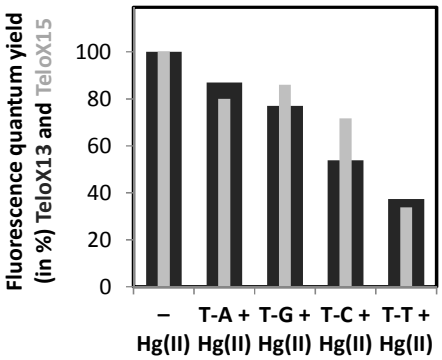


Figure 6.23. Fluorescence quantum yield changes of duplex DNA containing ^{DMA}**T** paired to **A**, **G** and **C** in **TeloX13** (black) and **TeloX15** (grey) upon addition of Hg^{II} ions (1.0 equiv.). All samples contained 4 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM $NaNO_3$) at pH = 7.35.

6.6. Conclusions

Fluorescent nucleosides represent a useful class of compounds that have revealed important properties of nucleic acids. To further develop this toolbox, ^{DMA}**T** was developed as fluorescent mimic of thymidine. Due to its effective mimicry of thymidine, ^{DMA}**T** exhibited very little impact on double-stranded DNA structure or stability. ^{DMA}**T** exhibited highly sensitive photophysical properties, such as blue-shifted excitation and emission spectra (λ_{abs} = 355 – 365 nm and λ_{em} = 486 – 505 nm) fluorescence anisotropy (r = 0.03 – 0.18) and fluorescence quantum yield (ϕ = 0.03 – 0.20), in the context of folded structures. These fluorescent parameters provide well-resolved readouts for monitoring DNA dynamics in real time

Mercury(II) is a highly toxic heavy metal that binds strongly to nucleic acids. In addition to probe duplex DNA containing Watson-Crick base pair, ^{DMA}**T**'s fluorescence properties were highly sensitive to the coordination of Hg^{II} ions to thymidine mismatches. Among many metals ions evaluated, ^{DMA}**T**'s fluorescence showed high selectivity for Hg^{II} ions and for the site-selective formation of T–Hg^{II}–T base pairs in the presence of other base pairing mismatches. ^{DMA}**T** therefore provides a new internal probe that can report the binding of Hg^{II} ions to thymidine in duplex DNA according to dramatic changes in fluorescence quantum yields (ϕ = 0.01 – 0.13). We believe that ^{DMA}**T** could be used to probe the impact of T–Hg^{II}–T base pairs in DNA conformational dynamics.

Chapter 7 | Experimental Section

Chapter 7

Experimental Section

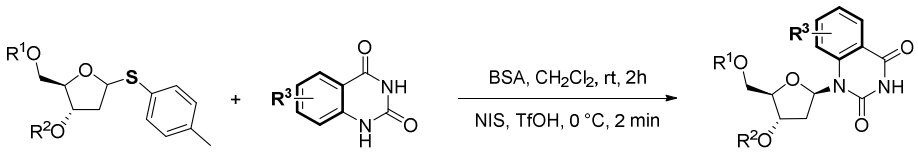
7.1. Synthetic Procedures and Characterization

7.1.1. General Experimental Methods

Starting materials were obtained in the highest commercial grades and used without further purification. Commercially available 1,3,5-tri-*O*-acetyl-2-deoxy-D-ribose **10** was purified by flash column chromatography on silica gel (Hexane/EtOAc, 6:4) prior to use. All reactions sensitive to moisture and/or air were carried out under an atmosphere of argon in dry, freshly distilled solvents under anhydrous conditions using oven-dried glassware. Commercially available dichloromethane (CH_2Cl_2) and tetrahydrofuran (THF) were purified by a solvent purification system under an atmosphere of argon immediately prior to use. Commercially available anhydrous dioxane was used directly without further drying. Analytical thin-layer chromatography was performed on precoated 250 µm layer thickness silica gel 60 F_{254} plates. Visualization was performed by ultraviolet light and staining with a 15% H_2SO_4 solution in EtOH/H_2O. Flash column chromatography was performed using 40-63 µm silica gel using compressed air. ^{1}H NMR spectra were recorded on 400 and 500 MHz spectrometers; residual solvent peaks were used as internal standards: DMSO (quint, δ^H = 2.50 ppm), $CHCl_3$ (s, δ^H = 7.26 ppm). ^{13}C-NMR spectra were recorded on 400 and 500 MHz spectrometers; δ relative to DMSO (δ 40.5 ppm) or $CHCl_3$ (δ 77.23 ppm). Coupling constants (*J*) are reported in hertz (Hz). The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, quint = quintet, sext = sextet, m = multiplet, dd = doublet-doublet, ddd = doublet-doublet-doublet, dt = doublet-triplet, dq =doublet-quartet, br = broad. Mass spectra were obtained on a quadrupole ion trap instrument equipped with an atmospheric pressure ion (API) source. High-resolution electrospray mass spectra (HR-ESI MS) were recorded on a QTOF-MS instrument. Infrared spectra were recorded on a FT/IR-4100 spectrometer.

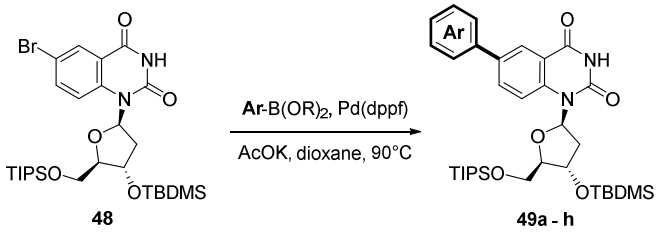
7.1.2. General Procedures

General Procedure A: *N*-2'-Deoxyribosylation of Thioglycoside Donor with Heterocyclic Acceptor



Procedure A: To a suspension of the heterocyclic acceptor (1.2 equiv) in CH_2Cl_2 (0.06 M) with activated molecular sieves (MS 4Å, two-fold mass excess as compared to the donor) was added BSA (2.5 equiv) dropwise over a 5 min period. The solution was stirring at room temperature for 2h, (complete dissolution was observed after 0.5 h). This solution was then cooled to 0 °C and the thioglycoside donor (1.0 equiv) in CH_2Cl_2 was added via cannula. After 5 min, NIS (1.2 equiv) and TMSOTf (0.6 equiv) or TfOH (0.4 equiv) were added (a deep red solution was formed) and the reaction mixture was stirred for 2 min at 0 °C. The reaction was then quenched with aq. sat. $Na_2S_2O_3$ and extracted with aq. sat. $NaHCO_3$ and CH_2Cl_2 (3 times each). The combined organic layers were dried over $MgSO_4$, filtered and evaporated *in vacuo* and subjected to column chromatography using silica gel (Hexane/EtOAc). Prior to glycosylation of heterocyclic acceptors **33 – 35** (1.2 equiv), each nucleobase was first heated in HMDS (4 mL) with a catalytic amount of $(NH_4)_2SO_4$ at 110 °C for 10h. The mixture was then evaporated, dried under vacuum and the residue used immediately.

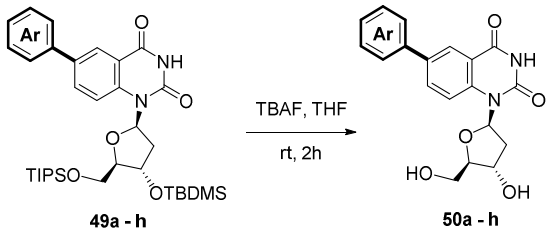
General Procedure B: Suzuki-Miyaura Cross-Couplings



Procedure B: In an oven-dried Schlenk tube backfilled with argon, a mixture of nucleoside **48** (1.0 equiv), boronic acid or ester (1.5 equiv), AcOK (1.5 equiv), $Pd(dppf)Cl_2$ (0.05 equiv), was heated at 90 °C in dioxane (0.08M) for 18h. The reaction was cooled to room temperature and diluted with CH_2Cl_2. The solution was filtered through a pad of Celite and extensively washed with CH_2Cl_2. The filtrate was

concentrated under reduced pressure, and the residual oil was purified by column chromatography on silica gel to provide biaryl nucleosides **49a – h**.

General Procedure C: Silyl Protecting Groups Removal

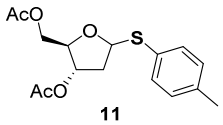


49a - h **50a - h**

Procedure C: To a stirred solution of the nucleoside **49a – h** (1.0 equiv) in THF, TBAF (4 equiv) dissolved in THF was added. The mixture was stirred at room temperature for 2h and quenched with sat sol NH_4Cl. The resulting solution was extracted with EtOAc (3 times), dried over $MgSO_4$ and evaporated in vacuo. The residue was purified by column chromatography on silica gel (CH_2Cl_2/MeOH). The obtained white solid was then washed with MeOH (2x), aqueous solution of HOAc (2x) and H_2O (1x) and finally dried under high vacuum to obtain nucleosides **50a – h**.

7.1.3. Procedures and Characterizations

p-Tolyl-3,5-di-*O*-acetyl-1-thio-2-deoxy-α,β-D-ribofuranoside (11)

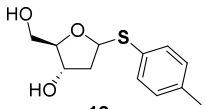


11

Procedure: To a stirring solution of 1,3,5-tri-*O*-acetyl-2-deoxy-α,β-D-ribose **10** (30.34 g, 116.58 mmoles) at -78 °C in CH_2Cl_2 (420 mL) was added *p*-TolSH (14.77 g, 118.91 mmoles). After 5 min, $BF_3{\cdot}OEt_2$ (50.36 mL, 408.03 mmoles) was added dropwise. The reaction mixture was stirred for 20 min and then quenched with aq. sat. $NaHCO_3$. The resulting solution was extracted with CH_2Cl_2. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 8:2) to give an inseparable isomer mixture of thioglycoside **11** (α/β = 1.0:1.8, 34.78 g, 92%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 7:3): 0.36; **^{1}H NMR** (2D COSY, 500 MHz, $CDCl_3$) **11α**: δ: 7.36-7.33 (m, 2H), 7.07-7.05 (m, 2H), 5.43 (dd, J_1 = 8.9 Hz, J_2 = 5.9 Hz, 1H, H_1), 5.09 (dt, J_1 = 6.1 Hz, J_2 = 1.7 Hz, 1H, H_3), 4.20-4.06 (m, 3H, H_5, H_5 and H_4), 2.31 (ddd, J_1 = 14.3, J_2 = 6.0 Hz, J_3 = 1.9 Hz, 1H, H_2), 2.27 (s, 3H), 2.20-2.14 (m, 1H, H_2), 2.00 (s, 3H), 1.98 (s, 3H); **11β**: δ: 7.36-7.33 (m, 2H), 7.07-7.05 (m, 2H), 5.60 (dd, J_1 = 7.8 Hz, J_2 = 3.0 Hz, 1H, H_1), 5.00 (ddd, J_1 = 7.6 Hz, J_2 = 4.4 Hz, J_3 = 3.0 Hz, 1H, H_3), 4.41 (q, J_1 = 4.5 Hz, 1H, H_4), 4.25 (dd, J_1 = 12.0 Hz, J_2 = 3.6 Hz, 1H, H_5), 4.20-4.14 (m, 1H, H_5), 2.73 (dt, J_1 = 15.5 Hz, J_2 = 7.8 Hz, 1H, H_2), 2.26 (s, 3H), 2.08-2.04 (m, 1H, H_2), 2.04 (s, 3H), 2.01 (s, 3H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 170.4, 170.3, 170.2, 170.1, 137.6, 137.0, 132.7, 131.5, 131.3, 129.4, 87.5, 86.0, 82.8, 80.1, 75.3, 73.7, 63.7, 63.2, 39.0, 37.9, 20.91, 20.88, 20.77, 20.74, 20.61, 20.56; **IR** (neat) ν: 2952, 1737, 1493, 1366, 1222, 1048, 1018, 809 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{16}H_{20}O_5SNa$, 347.09291; found 347.09245.

p-Tolyl-1-thio-2-deoxy-α,β-D-ribofuranoside (12)



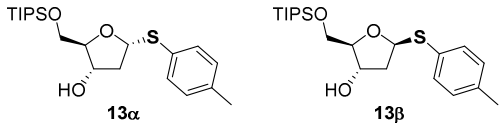
12

Procedure: To a stirring solution of thioglycoside **11** (α/β = 1.0:1.8, 20.0 g, 61.65 mmoles) in a mixture of MeOH/CH_2Cl_2 (202.5 mL, 4:1) was added K_2CO_3 (18.74 g, 135.64 mmoles). The reaction mixture was stirred at room temperature for 2h and then quenched with a solution of HCl (500 mL, 1N). The mixture was extracted with $CHCl_3$ (3 x 400 mL), dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5) to give an inseparable isomer mixture of thioglycoside **12** (α/β = 1.0:1.8, 14.46 g, 98%) as a colorless oil.

Characterization: R_f (CH_2Cl_2/MeOH, 94:6): 0.25; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **12α**: δ: 7.41-7.39 (m, 2H), 7.14-7.12 (m, 2H), 5.64-5.61 (m, 1H, H_1), 4.40-4.37 (m, 1H, H_3), 4.00 (q, J_1 = 4.0 Hz, 1H, H_4), 3.70-3.68 (m, 1H, H_5), 3.60 (dd, J_1 = 12.0 Hz, J_2 = 4.3 Hz, 1H, H_5), 2.37-2.34 (m, 1H, H_2), 2.29 (s, 3H), 2.30 (s, 2H), 2.40-2.17 (m, 1H, H_2); **12β**: δ: 7.41-7.39 (m, 2H), 7.14-7.12 (m, 2H), 5.63 (dd, J_1 = 7.4 Hz, J_2 = 3.4 Hz, 1H, H_1), 4.27 (ddd, J_1 = 7.5 Hz, J_2 = 4.6 Hz, J_3 = 3.8 Hz, 1H, H_3), 4.16 (q, J_1 = 4.0 Hz, 1H, H_4), 3.78 (dd, J_1 = 12.0 Hz, J_2 = 3.7 Hz, 1H, H_5), 3.70 (dd, J_1 = 12.0 Hz, J_2 = 4.0 Hz, 1H, H_5), 2.64 (dt, J_1 = 14.2 Hz, J_2 = 7.4 Hz, 1H, H_2), 2.33 (s, 3H), 2.30 (s, 2H), 2.04 (dt, J_1 = 13.8 Hz, J_2 = 3.5 Hz, 1H, H_2); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 138.0, 132.6, 132.3, 130.9, 130.0, 129.9, 88.3, 87.3, 86.5, 85.6,

72.8, 72.3, 62.9, 62.4, 42.1, 41.9, 21.32, 21.29; **IR** (neat) ν: 3386, 2921, 2872, 1492, 1079, 1038, 808, 500 cm^{-1}; **HR-ESI MS** (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}O_3SNa$, 263.07179; found 263.07125.

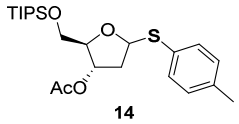
***p*-Tolyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-α-D-ribofuranoside (13α)** and ***p*-Tolyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-β-D-ribofuranoside (13β)**



Procedure: To a stirring solution of thioglycoside **12** (α/β = 1.0:1.8, 14.0 g, 58.26 mmoles) in CH_2Cl_2 (320 mL) were added TIPSCl (14.96 mL, 69.91 mmoles) and imidazole (4.82 g, 69.91 mmoles). The reaction mixture was stirred at room temperature for 16h and then quenched with H_2O (400 mL). The resulting solution was extracted with CH_2Cl_2 (2 x 300 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (CH_2Cl_2/MeOH, 99:1) to give the separable isomers of thioglycoside **13α** and **13β** (19.64 g, 85%) as colorless oils.

Characterization: R_f (CH_2Cl_2/MeOH, 96:4): 0.63 (**13α**) and 0.44 (**13β**); ^1H NMR (2D COSY, 400 MHz, $CDCl_3$) **13α**: δ: 7.39 (d, J_1 = 8.0 Hz, 2H), 7.10 (d, J_1 = 7.8 Hz, 2H), 5.55 (t, J_1 = 7.0 Hz, 1H, H_1), 4.39 (quint, J_1 = 3.4 Hz, 1H, H_3), 3.95-3.01 (m, 1H, H_4), 3.85 (dd, J_1 = 9.9 Hz, J_2 = 4.7 Hz, 1H, H_5), 3.50 (t, J_1 = 9.3 Hz, 1H, H_5), 2.30 (s, 3H), 2.35-2.29 (m, 1H, H_2), 2.23-2.16 (m, 1H, H_2), 2.07 (s, br, 1H), 1.10-1.05 (m, 21H); **13β**: δ: 7.40 (d, J_1 = 7.4 Hz, 2H), 7.11 (d, J_1 = 7.7 Hz, 2H), 5.64 (dd, J_1 = 7.3 Hz, J_2 = 3.4 Hz, 1H, H_1), 4.34-4.31 (m, 1H, H_3), 4.16 (q, J_1 = 4.2 Hz, 1H, H_4), 3.93 (dd, J_1 = 10.1 Hz, J_2 = 3.5 Hz, 1H, H_5), 3.70 (dd, J_1 = 10.1 Hz, J_2 = 5.8 Hz, 1H, H_5), 2.73-2.65 (m, 1H, H_2), 2.33 (s, br, 1H), 2.33 (s, 3H), 2.05 (dt, J_1 = 13.9 Hz, J_2 = 3.3 Hz, 1H, H_2), 1.08-1.05 (m, 21H); ^{13}C NMR (100 MHz, $CDCl_3$) **13α**: δ: 137.6, 132.7, 130.5, 129.7, 87.4, 86.1, 74.0, 64.7, 40.5, 21.2, 18.13, 18.11, 12.0; **13β**: δ: 137.6, 132.2, 131.3, 129.8, 87.1, 85.8, 73.9, 64.4, 41.6, 21.3, 18.15, 18.12, 12.08, 12.06; IR (neat) **13α**: ν: 3419, 2940, 2923, 2864, 1492, 1461, 1129, 1066, 996, 881, 807, 773, 680, 503; **13β**: ν: 3419, 2940, 2923, 2864, 1492, 1461, 1127, 1062, 881, 806, 773, 681, 501 cm^{-1}; HR-ESI MS (*m/z*): $[M+Na]^+$ calcd for $C_{21}H_{36}O_3SSiNa$, 419.20521; found **13α**: 419.20493; **13β**: 419.20442.

p-Tolyl-3-*O*-acetyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-α,β-D-ribofuranoside (14)

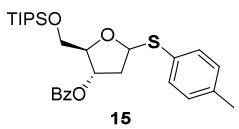


14

Procedure: To a stirring solution of thioglycosides **13α** and **13β** (α/β = 1.0:1.8, 2.05 g, 5.17 mmoles) in MeCN (35 mL), were added Et_3N (0.86 mL, 6.20 mmoles), Ac_2O (0.58 mL, 6.20 mmoles) and DMAP (0.063 g, 0.52 mmoles). The reaction mixture was stirred at room temperature for 2h and then quenched with H_2O (50 mL). The resulting solution was extracted with CH_2Cl_2 (2 x 50 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 93:7) to give an inseparable isomer mixture of thioglycoside **14** (α/β = 1.0:1.8, 2.04 g, 90%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 9:1): 0.38; **^{1}H NMR** (2D COSY, 500 MHz, $CDCl_3$) **14α**: δ: 7.41-7.39 (m, 2H), 7.12-7.10 (m, 2H), 5.47 (dd, J_1 = 7.7, J_2 = 3.5 Hz, 1H, H_1), 5.31-5.29 (m, 1H, H_3), 4.08 (t, J_1 = 5.1 Hz, 1H, H_4), 3.84-3.81 (m, 1H, H_5), 3.58 (dd, J_1 = 10.5 Hz, J_2 = 6.6 Hz, 1H, H_5), 2.33-2.31 (m, 1H, H_2), 2.32 (s, 3H), 2.26-2.21 (m, 1H, H_2), 2.04 (s, 3H), 1.07-1.04 (m, 21H); **14β**: δ: 7.41-7.39 (m, 2H), 7.12-7.10 (m, 2H), 5.68 (dd, J_1 = 7.7 Hz, J_2 = 3.0 Hz, 1H, H_1), 5.24-5.22 (m, 1H, H_3), 4.32 (q, J_1 = 3.3 Hz, 1H, H_4), 3.93 (dd, J_1 = 10.8 Hz, J_2 = 3.3 Hz, 1H, H_5), 3.83 (dd, J_1 = 14.4 Hz, J_2 = 3.7 Hz, 1H, H_5), 2.81-2.77 (m, 1H, H_2), 2.32 (s, 3H), 2.11-2.09 (m, 1H, H_2), 2.09 (s, 3H), 1.07-1.04 (m, 21H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 170.8, 170.3, 137.6, 136.9, 132.7, 132.3, 131.5, 130.5, 129.7, 129.6, 88.1, 86.4, 86.0, 84.4, 76.0, 74.9, 63.99, 63.91, 39.9, 38.4, 21.20, 21.16, 21.14, 18.10, 18.09, 18.05, 18.03, 12.0; **IR** (neat) ν: 2942, 2892, 2866, 1741, 1240, 1065, 1017, 882, 773 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{23}H_{38}O_4SSiNa$, 461.21578; found 461.21509.

p-Tolyl-3-*O*-benzoyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-α,β-D-ribofuranoside (15)



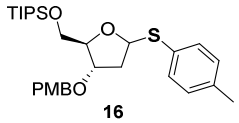
15

Procedure: To a stirring solution of thioglycoside **13** (α/β = 1.0:1.8, 0.405 g, 1.02 mmoles) in CH_2Cl_2 (7.3 mL), were added pyridine (0.80 mL), Et_3N (0.20 mL, 1.43 mmoles), BzCl (0.14 mL, 1.23 mmoles) and DMAP (0.012 g, 0.10 mmoles). The reaction mixture was stirred at room temperature for

2h and then quenched with a solution of HCl (30 mL, 1N). The resulting solution was extracted with CH_2Cl_2 (2 x 30 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 95:5) to give an inseparable isomer mixture of thioglycoside **15** (α/β = 1.0:1.8, 0.445 g, 87%) as a colorless oil.

Characterization: $\mathbf{R_f}$ (Hexane/EtOAc, 9:1): 0.50; **^{1}H NMR** (2D COSY, 500 MHz, $CDCl_3$) **15α**: δ: 8.02 (d, J_1 = 8.0 Hz, 2H), 7.59-7.54 (m, 1H), 7.47-7.43 (m, 4H), 7.13-7.10 (m, 2H), 5.57-5.53 (m, 2H, H_1 and H_3), 4.26 (t, J_1 = 5.7 Hz, 1H, H_4), 3.92-3.90 (m, 1H, H_5), 3.67 (dd, J_1 = 10.1 Hz, J_2 = 6.4 Hz, 1H, H_5), 2.47 (dd, J_1 = 14.2 Hz, J_2 = 5.7 Hz, 1H, H_2), 2.41-2.37 (m, 1H, H_2), 2.34 (s, 3H), 1.08-1.05 (m, 21H); **15β**: δ: 8.12 (d, J_1 = 18.0 Hz, 2H), 7.59-7.54 (m, 1H), 7.47-7.43 (m, 4H), 7.13-7.10 (m, 2H), 5.78 (d, J_1 = 5.8 Hz, 1H, H_1), 5.57-5.53 (m, 1H, H_3), 4.49 (d, J_1 = 2.9 Hz, 1H, H_4), 4.03 (dd, J_1 = 10.9 Hz, J_2 = 3.0 Hz, 1H, H_5), 3.91 (dd, J_1 = 11.1 Hz, J_2 = 2.4 Hz, 1H, H_5), 2.90-2.84 (m, 1H, H_2), 2.32 (s, 3H), 2.32-2.30 (m, 1H, H_2), 1.08-1.05 (m, 21H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 166.5, 166.0, 137.8, 137.1, 133.34, 133.32, 132.9, 132.5, 131.7, 130.6, 130.2, 130.1, 130.0, 129.84, 129.82, 129.81, 129.80, 128.6, 128.5, 88.7, 86.7, 86.2, 85.1, 76.7, 75.7, 64.15, 64.14, 40.2, 38.8, 21.3, 21.2, 18.22, 18.21, 18.16, 18.14, 12.1; **IR** (neat) ν: 2942, 2892, 2866, 1719, 1271, 1110, 1066, 882, 773, 711 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{28}H_{40}O_4SSiNa$, 523.23143; found 523.23068.

p-Tolyl-3-*O*-*p*-methoxybenzoyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-α,β-D-ribofuranoside (16)

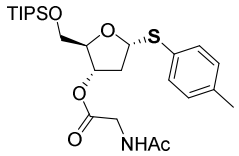


16

Procedure: To a stirring solution of thioglycoside **13** (α/β = 1.0:1.8, 0.30 g, 0.76 mmoles) in CH_2Cl_2 (4.0 mL), were added pyridine (0.4 mL), Et_3N (157 μL, 1.13 mmoles), *p*-anisoyl chloride (PMBCl, 0.193 g, 1.13 mmoles) and DMAP (0.018 g, 0.151 mmoles). The reaction mixture was stirred at room temperature for 16h and then quenched with a solution of HCl (20 mL, 0.5N). The resulting solution was extracted with CH_2Cl_2 (2 x 20 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 9:1) to give an inseparable isomer mixture of thioglycoside **16** (α/β = 1.0:1.8, 0.313 g, 78%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 9:1): 0.35; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **16α**: δ: 7.98 (d, J_1 = 9.0 Hz, 2H), 7.46 (d, J_1 = 8.0 Hz, 2H), 7.14-7.10 (m, 2H), 6.92-6.89 (m, 2H), 5.57 (dd, J_1 = 9.5 Hz, J_2 = 5.7 Hz, 1H, H_1), 5.52-5.49 (m, 1H, H_3), 4.27-4.25 (m, 1H, H_4), 3.93-3.88 (m, 1H, H_5), 3.84 (s, 3H), 3.68 (dd, J_1 = 10.5 Hz, J_2 = 6.2 Hz, 1H, H_5), 2.49 (dd, J_1 = 14.2 Hz, J_2 = 5.7 Hz, 1H, H_2), 2.39 (dd, J_1 = 9.5 Hz, J_2 = 5.8 Hz, 1H, H_2), 2.32 (s, 3H), 1.10-1.05 (m, 21H); **16β**: δ: 8.08 (d, J_1 = 8.5 Hz, 2H), 7.45 (d, J_1 = 7.8 Hz, 2H), 7.11 (d, J_1 = 7.8 Hz, 2H), 6.94 (d, J_1 = 8.5 Hz, 2H), 5.78 (dd, J_1 = 7.7 Hz, J_2 = 2.2 Hz, 1H, H_1), 5.51 (d, J_1 = 7.1 Hz, 1H, H_3), 4.48 (d, J_1 = 2.9 Hz, 1H, H_4), 4.03 (dd, J_1 = 10.9 Hz, J_2 = 3.1 Hz, 1H, H_5), 3.91 (dd, J_1 = 10.9 Hz, J_2 = 3.1 Hz, 1H, H_5), 3.86 (s, 3H), 2.86 (quint, J_1 = 7.2 Hz, 1H, H_2), 2.32 (s, 3H), 2.30-2.27 (m, 1H, H_2), 1.10-1.05 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 166.1, 165.7, 163.6, 137.7, 136.9, 132.8, 132.5, 132.0, 131.8, 131.5, 130.6, 129.75, 129.72, 122.5, 122.4, 113.8, 113.7, 88.7, 86.6, 86.2, 85.2, 76.3, 75.3, 64.12, 64.10, 55.5, 40.2, 38.7, 21.25, 21.21, 18.17, 18.15, 18.11, 18.08, 12.0; **IR** (neat) ν: 2937, 2865, 1712, 1605, 1273, 1254, 1166, 1098, 881, 769 cm^{-1}; **HR-ESI MS** (m/z): $[M+Na]^+$ calcd for $C_{29}H_{42}O_5SSiNa$, 553.24199; found 553.24099.

***p*-Tolyl-3-*O*-(*N*-acetyl)-glycyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-α-D-ribofuranoside (17α)**



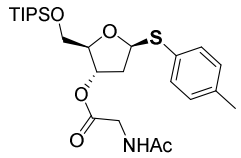
17α

Procedure: To a stirring solution of thioglycoside **13α** (0.50 g, 1.26 mmoles) in CH_2Cl_2 (8.0 mL), were added *N*-acetylglycine (0.177 g, 1.51 mmoles), DCI (0.24 mL, 1.51 mmoles) and DMAP (0.01 g, 0.078 mmoles). The reaction mixture was stirred at room temperature for 12h and then quenched with a sat. sol. $NaHCO_3$ (40 mL). The resulting solution was extracted with CH_2Cl_2 (2 x 40 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The residue was taken up in a mixture of Hexane/EtOAc (5 mL, 10:1) and the precipitate was filtered off through cotton in a glass pipette. The resulting solution was evaporated and the crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 1:1) to give thioglycoside **17α** (0.565 g, 91%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 1:1): 0.23; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) δ: 7.37 (d, J_1 = 8.2 Hz, 2H), 7.09 (d, J_1 = 7.9 Hz, 2H), 6.29 (t, br, 1H), 5.43 (dd, J_1 = 9.4 Hz, J_2 = 5.8 Hz, 1H, H_1),

5.35 (d, J_1 = 5.6 Hz, 1H, H_3), 4.08-4.03 (m, 1H, H_4), 3.97 (ddd, AB system, J_1 = 27.5 Hz, J_2 = 18.3 Hz, J_3 = 5.2 Hz, 2H), 3.79 (dd, J_1 = 10.6 Hz, J_2 = 4.2 Hz, 1H, H_5), 3.53 (dd, J_1 = 10.6 Hz, J_2 = 6.6 Hz, 1H, H_5), 2.36-2.32 (m, 1H, H_2), 2.29 (s, 3H), 2.26-2.19 (m, 1H, H_2), 1.99 (s, 3H), 1.05-1.03 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 170.4, 169.6, 137.8, 132.8, 130.1, 129.7, 86.3, 85.7, 77.4, 63.7, 41.5, 38.2, 23.5, 22.9, 18.07, 18.06, 11.9; **IR** (neat) ν: 3333, 2941, 2891, 2865, 1750, 1656, 1541, 1463, 1381, 1193, 1064, 882, 683 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{25}H_{41}NO_5SSiNa$, 518.23724; found 518.23653.

***p*-Tolyl-3-*O*-(*N*-acetyl)-glycyl-1-thio-5-*O*-triisopropylsilyl-2-deoxy-β-D-ribofuranoside (17β)**

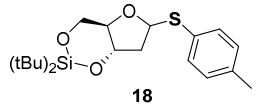


17β

Procedure: To a stirring solution of **13β** (0.50 g, 1.26 mmoles) in CH_2Cl_2 (8.0 mL), were added *N*-acetylglycine (0.177 g, 1.51 mmoles), DCI (0.24 mL, 1.51 mmoles) and DMAP (0.01 g, 0.078 mmoles). The reaction mixture was stirred at room temperature for 12h and then quenched with a sat. sol. $NaHCO_3$ (40 mL). The resulting solution was extracted with CH_2Cl_2 (2 x 40 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The residue was taken up in a mixture of Hexane/EtOAc (5mL, 10:1) and the precipitate was filtered off through cotton in a glass pipette. The resulting solution was evaporated and the crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 1:1) to give thioglycoside **17β** (0.53 g, 85%) as a colorless oil.

Characterization: **R_f** (Hexane/EtOAc, 5:5): 0.25; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) δ: 7.38 (d, J_1 = 8.2 Hz, 2H), 7.10 (d, J_1 = 8.0 Hz, 2H), 6.07 (s, br, 1H), 5.69 (dd, J_1 = 7.7 Hz, J_2 = 2.7 Hz, 1H, H_1), 5.33 (dt, J_1 = 6.8 Hz, J_2 = 2.2 Hz, 1H, H_3), 4.32 (q, J_1 = 3.1 Hz, 1H, H_4), 4.07 (dd, J_1 = 5.3 Hz, J_2 = 1.9 Hz, 2H), 3.92 (dd, J_1 = 10.9 Hz, J_2 = 3.2 Hz, 1H, H_5), 3.82 (dd, J_1 = 10.9 Hz, J_2 = 3.4 Hz, 1H, H_5), 2.77 (quint, J_1 = 7.5 Hz, 1H, H_2), 2.32 (s, 3H), 2.14 (dt, J_1 = 14.6 Hz, J_2 = 2.4 Hz, 1H, H_2), 2.04 (s, 3H), 1.07-1.04 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 170.3, 169.8, 137.3, 131.9, 131.7, 129.8, 88.3, 84.7, 76.3, 63.8, 41.8, 39.8, 23.1, 21.2, 18.1, 18.0, 12.0; **IR** (neat) ν: 3333, 2942, 2866, 1750, 1659, 1549, 1493, 1383, 1193, 1064, 882, 679 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{25}H_{41}NO_5SSiNa$, 518.23724; found 518.23681.

p-Tolyl-3,5-*O*-(di-*tert*-butylsilylene)-1-thio-2-deoxy-α,β-D-ribofuranoside (18)

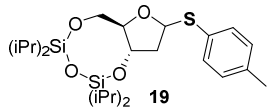


18

Procedure: To a stirring solution of thioglycoside **12** (α/β = 1.0:1.8, 0.150 g, 0.62 mmoles) in a mixture of CH_2Cl_2/DMF (6 mL, 5:1) at 0 °C, was added di-*tert*-butylsilylbistrifluoromethanesulfonate (tBu_2SiOTf_2, 0.24 mL, 0.75 mmoles) and 2,6-lutidine (0.30 g, 2.81 mmoles). The reaction mixture was stirred at room temperature for 16h and then MeOH was added. The resulting solution was evaporated *in vacuo* and the residue was taken in CH_2Cl_2 (10 mL) and washed with H_2O and brine. The organic layer was dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 98:2) to give an inseparable isomer mixture of thioglycoside **18** (α/β = 1.0:1.8, 0.197 g, 83%) as a white solid.

Characterization: $\mathbf{R_f}$ (Hexane/EtOAc, 98:2): 0.22; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **18α**: δ: 7.43-7.40 (m, 2H), 7.15-7.11 (m, 2H), 5.46 (dd, J_1 = 8.6 Hz, J_2 = 2.6 Hz, 1H, H_1), 4.33 (dd, J_1 = 9.3 Hz, J_2 = 4.8 Hz, 1H, H_4), 3.84-3.55 (m, 3H, H_3, H_5 and H_5), 2.42-2.34 (m, 2H, H_2 and H_2), 2.33 (s, 3H), 1.06 (s, 6H), 1.01 (s, 6H), 0.96 (s, 6H); **18β**: δ: 7.43-7.40 (m, 2H), 7.15-7.11 (m, 2H), 5.51 (dd, J_1 = 8.0 Hz, J_2 = 7.1 Hz, 1H, H_1), 4.38 (dd, J_1 = 9.0 Hz, J_2 = 4.8 Hz, 1H, H_4), 4.04-3.98 (m, 1H, H_3), 3.92-3.82 (m, 2H, H_5 and H_5), 2.81 (dt, J_1 = 12.7 Hz, J_2 = 7.1 Hz, 1H, H_2), 2.32 (s, 3H), 1.87 (ddd, J_1 = 12.6 Hz, J_2 = 9.8 Hz, J_3 = 8.1 Hz, 1H, H_2), 1.06 (s, 6H), 1.01 (s, 6H), 0.96 (s, 6H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 138.7, 137.6, 134.3, 131.9, 131.5, 129.84, 129.81, 129.0, 85.6, 84.7, 79.0, 75.7, 75.2, 74.9, 68.2, 67.7, 39.7, 39.2, 27.6, 27.5, 27.34, 27.33, 22.8, 22.6, 21.32, 21.28, 20.29, 20.20; **IR** (neat) ν: 2962, 2933, 2887, 2859, 1473, 1127, 1065, 1050, 900, 827, 808, 753, 652, 418 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{20}H_{32}O_3SSiNa$, 403.17391; found 403.17338.

p-Tolyl-3,5-*O*-(tetraisopropylsiloxane-1,3-diyl)-1-thio-2-deoxy-α,β-D-ribofuranoside (19)



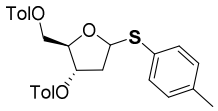
19

Procedure: To a stirring solution of thioglycoside **12** (α/β = 1.0:1.8, 0.150 g, 0.62 mmoles) in pyridine (0.8 mL) at 0 °C, was added 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (TIPDSCl, 0.21 mL,

0.655 mmoles). The reaction mixture was stirred at room temperature for 12h and then MeOH was added. The resulting solution was diluted with CH_2Cl_2 (10 mL) and washed with H_2O and brine. The organic layer was dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 98:2) to give an inseparable isomer mixture of thioglycoside **19** (α/β = 1.0:1.8, 0.279 g, 93%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 96:4): 0.34; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **19α**: δ: 7.42-7.39 (m, 2H), 7.11-7.08 (m, 2H), 5.47 (dd, J_1 = 7.2 Hz, J_2 = 3.8 Hz, 1H, H_1), 4.41 (q, J_1 = 5.7 Hz, 1H, H_3), 4.07-3.90 (m, 2H, H_5 and H_5), 3.88-3.85 (m, 1H, H_4), 2.47-2.41 (m, 1H, H_2), 2.40-2.34 (m, 1H, H_2), 2.34 (s, 3H), 1.11-1.04 (s, 32H); **19β**: δ: 7.42-7.39 (m, 2H), 7.11-7.08 (m, 2H), 5.53 (t, J_1 = 6.6 Hz, 1H, H_1), 4.31 (q, J_1 = 7.2 Hz, 1H, H_3), 4.00 (dd, J_1 = 11.8 Hz, J_2 = 2.8 Hz, 1H, H_5), 3.99-3.91 (m, 2H, H_4 and H_5), 2.74 (dt, J_1 = 15.4 Hz, J_2 = 7.6 Hz, 1H, H_2), 2.33 (s, 3H), 2.07-2.00 (m, 1H, H_2), 1.11-1.04 (s, 32H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 137.9, 137.0, 133.1, 132.3, 131.53, 131.52, 130.2, 129.79, 129.70, 129.69, 85.9, 85.7, 85.5, 81.7, 74.3, 71.1, 65.4, 61.8, 42.0, 40.6, 21.27, 21.23, 17.74, 17.67, 17.65, 17.57, 17.56, 17.54, 17.55, 17.54, 17.50, 17.48, 17.46, 17.44, 17.32, 17.31, 17.22, 17.16, 17.15, 17.13, 13.6, 13.5, 13.3, 13.1; **IR** (neat) ν: 2944, 2892, 2867, 1464, 1137, 1081, 1054, 1035, 999, 885, 776, 692 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{24}H_{42}O_4SSi_2Na$, 505.22400; found 505.22353.

***p*-Tolyl-3,5-di-O-p-toluoyl-1-thio-2-deoxy-α,β-D-ribofuranoside (20)**

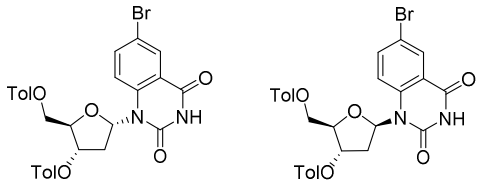


20

Procedure: To a stirring solution of thioglycoside **12** (α/β = 1.0:1.8, 0.150 g, 0.62 mmoles) in CH_2Cl_2 (4.5 mL) at 0 °C, were added pyridine (0.5 mL), Et_3N (0.19 mL, 1.37 mmoles), DMAP (0.031 g, 0.25 mmoles), and *p*-toluoyl chloride (0.18 mL, 1.37 mmoles). The reaction mixture was stirred at room temperature for 16h and the color of the solution changed from orange to green. The resulting solution was diluted with CH_2Cl_2 (10 mL) and washed with HCl (10 mL, 1N) and brine. The organic layer was dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 95:5 → 90:10) to give an inseparable isomer mixture of thioglycoside **20** (α/β = 1.0:1.8, 0.255 g, 86%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 9:1): 0.16; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **20α**: δ: 8.05-7.92 (m, 4H), 7.49-7.45 (m, 2H), 7.27-7.22 (m, 4H), 7.13-7.09 (m, 2H), 5.64 (dd, J_1 = 8.8 Hz, J_2 = 5.9 Hz, 1H, H_1), 5.53-5.49 (m, 1H, H_3), 4.56-4.48 (m, 3H, H_4, H_5 and H_5), 2.57-2.40 (m, 2H, H_2 and H_2), 2.42-2.40 (m, 6H), 2.29 (s, 3H); **20β**: δ: 8.05-7.92 (m, 4H), 7.49-7.45 (m, 2H), 7.27-7.22 (m, 4H), 7.13-7.09 (m, 2H), 5.80 (dd, J_1 = 7.6 Hz, J_2 = 2.5 Hz, 1H, H_1), 5.53-5.49 (m, 1H, H_3), 4.79 (q, J_1 = 4.0 Hz, 1H, H_4), 4.65-4.61 (m, 2H, H_5 and H_5), 2.98-2.91 (m, 1H, H_2), 2.40-2.33 (m, 1H, H_2), 2.42-2.40 (m, 6H), 2.33 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 166.2, 165.9, 144.2, 144.1, 143.8, 143.7, 138.0, 137.7, 137.4, 133.3, 132.23, 132.02, 132.01, 132.00, 131.7, 129.9, 129.87, 129.85, 129.84, 129.77, 129.76, 129.75, 129.4, 129.25, 129.23, 129.19, 129.17, 127.17, 127.15, 126.9, 126.8, 88.3, 86.4, 83.2, 81.3, 76.0, 74.80, 64.4, 64.0, 39.6, 38.5, 21.76, 21.74, 21.72, 21.16; **IR** (neat) ν: 2358, 2338, 1716, 1608, 1268, 1177, 1104, 906, 752, 726 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{28}H_{28}O_5SNa$, 499.15552; found 499.15490.

1'-(3',5'-Di-*O*-*p*-toluoyl-2'-deoxy-α-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (23α) and **1'-(3',5'-di-*O*-*p*-toluoyl-2'-deoxy-β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (23β)**



According to the general procedure A: (Starting from thioglycoside **20**): The isomer mixture of nucleosides **23α** and **23β** (α/β = 1.0:1.6, 0.390 g, 72%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 6:4) using 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21** (0.264 g, 1.10 mmoles), BSA (0.56 mL, 2.28 mmoles), thioglycoside **20** (α/β = 1.0:1.8, 0.435 g, 0.91 mmoles), NIS (0.247 g, 1.10 mmoles) and TMSOTf (100 μL, 0.55 mmoles). The β-nucleoside **23β** was separated from **23α** via precipitation in EtOAc (two precipitations allowed complete removal of the α-nucleoside).

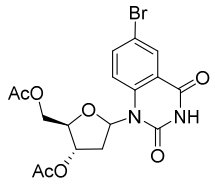
Characterization: 23β: R_f (Hexane/EtOAc, 6:4): 0.26; **^{1}H NMR** (2D COSY and ROESY, 500 MHz, $CDCl_3$) δ: 9.14 (s, 1H), 8.25 (d, J_1 = 2.5 Hz, 1H), 7.95 (d, J_1 = 1.5 Hz, 2H), 7.94 (d, J_1 = 1.5 Hz, 2H), 7.52 (d, J_1 = 9.0 Hz, 1H), 7.28-7.25 (m, 4H), 6.97 (dd, J_1 = 8.9 Hz, J_2 = 2.5 Hz, 1H), 6.86 (t, J_1 = 8.2 Hz, 1H, $H_{1'}$), 5.81 (ddd, J_1 = 8.2 Hz, J_2 = 4.8 Hz, J_3 = 3.2 Hz, 1H, $H_{3'}$), 4.91 (dd, J_1 = 12.2 Hz, J_2 = 2.9 Hz, 1H, $H_{5'}$),

4.66 (dd, J_1 = 12.3 Hz, J_2 = 4.0 Hz, 1H, $H_{5'}$), 4.45 (q, J_1 = 4.0 Hz, 1H, $H_{4'}$), 3.12 (dt, J_1 = 17.0 Hz, J_2 = 8.5 Hz, 1H, $H_{2'}$), 2.44 (s, 3H), 2.41 (s, 3H), 2.41-2.38 (m, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 166.3, 166.2, 160.5, 149.8, 144.6, 144.5, 138.5, 138.0, 131.3, 130.0, 129.9, 129.5, 129.4, 127.1, 126.5, 118.5, 118.2, 117.1, 84.4, 81.6, 73.5, 63.3, 34.4, 21.94, 21.92; **IR** (neat) ν: 3218, 3073, 1703, 1603, 1483, 1465, 1311, 1266, 1177, 1093, 1019, 751, 729, 502 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{29}H_{25}BrN_2O_7Na$, 615.07428; found 615.07435.

Alternative procedure: (Starting from α-glycosyl chloride **22α**): To a suspension of 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21** (0.720 g, 0.175 mmoles) in CH_2Cl_2 (12 mL) with activated molecular sieves (MS 4Å, two-fold mass excess as compared to the donor) was added BSA (0.37 mL, 1.50 mmoles) dropwise over a 5 min period. The solution was stirring at room temperature for 2h, (the complete dissolution was observed after 0.5h). This solution was then cooled to 0 °C and the α-glycoside chloride **22α** (0.233 g, 0.60 mmoles) in CH_2Cl_2 was added. After 5 min, CuI (0.137 g, 0.72 mmoles) was added and the reaction mixture was stirred for 48h at room temperature. The solution was then filtered through Celite, washed with a sat. sol. NaCl and extracted with CH_2Cl_2 (3 times). The combined organics layers were dried over $MgSO_4$, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (CH_2Cl_2/MeOH, 98.5:1.5) to give the inseparable isomers of nucleosides **23α** and **23β** (α/β = 3.5:1.0, 0.206 g, 58%) as an off-white foam.

Characterization: R_f (Hexane/EtOAc, 6:4): 0.28; **^{1}H NMR** (2D COSY and ROESY, 400 MHz, $CDCl_3$) **23α**: δ: 9.38 (s, br, 1H), 8.32 (d, J_1 = 2.2 Hz, 1H), 7.96-7.94 (m, 4H), 7.68-7.62 (m, 2H), 7.29-7.23 (m, 4H), 7.02 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.62 (quint, J_1 = 4.5 Hz, 1H, $H_{3'}$), 4.91 (q, J_1 = 3.7 Hz, 1H, $H_{4'}$), 4.70 (dd, J_1 = 12.0 Hz, J_2 = 4.6 Hz, 1H, $H_{5'}$), 4.55 (dd, J_1 = 12.0 Hz, J_2 = 3.6 Hz, 1H, $H_{5'}$), 3.02 (dt, J_1 = 16.3 Hz, J_2 = 8.2 Hz, 1H, $H_{2'}$), 2.78 (ddd, J_1 = 14.5 Hz, J_2 = 7.6 Hz, J_3 = 4.6 Hz, 1H, $H_{2'}$), 2.44 (s, 3H), 2.38 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) **23α** and **23β**: δ: 166.4, 166.3, 166.2, 160.6, 160.5, 149.9, 149.8, 144.8, 144.6, 144.5, 144.2, 138.4, 138.3, 137.9, 137.8, 131.5, 131.3, 130.0, 129.92, 129.91, 129.90, 129.89, 129.87, 129.8, 129.5, 129.46, 129.44, 129.35, 129.32, 129.30, 129.2, 127.1, 126.8, 126.5, 118.6, 118.4, 118.19, 118.18, 117.1, 117.0, 85.6, 84.4, 81.8, 81.5, 75.2, 73.4, 64.9, 63.3, 34.8, 34.4, 21.92, 21.90, 21.89, 21.85; **IR** (neat) ν: 3208, 1705, 1604, 1483, 1465, 1362, 1310, 1267, 1177, 1094, 1020, 750, 503 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{29}H_{25}BrN_2O_7Na$, 615.07428; found 615.07444.

1'-(3',5'-Di-*O*-acetyl-2'-deoxy-α,β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (24α) and **(24β)**

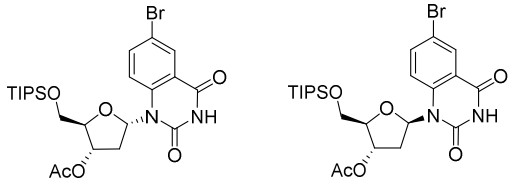


24

According to the general procedure A: The inseparable isomer mixture of nucleosides **24α** and **24β** (α/β = 1.2:1.0, 0.207 g, 76%) were obtained as an off-white foam after column chromatography on silica gel (Hexane/EtOAc, 1:1) using 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.178 g, 0.74 mmoles), BSA (0.38 mL, 1.54 mmoles), thioglycoside **11** (α/β = 1.0:1.8, 0.20 g, 0.62 mmoles), NIS (0.166 g, 0.74 mmoles) and TMSOTf (66 μL, 0.37 mmoles).

Characterization: $\mathbf{R}_f$ (Hexane/EtOAc, 4:6): 0.33; **^{1}H NMR** (2D COSY and ROESY, 400 MHz, $CDCl_3$) **24α**: δ: 9.67 (s, br, 1H), 8.32 (d, J_1 = 2.5 Hz, 1H), 7.67 (dd, J_1 = 9.0 Hz, J_2 = 2.5 Hz, 1H), 7.53 (d, J_1 = 9.0 Hz, 1H), 6.77 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.31-5.27 (m, 1H, $H_{3'}$), 4.62 (q, J_1 = 4.6 Hz, 1H, $H_{4'}$), 4.32 (dd, J_1 = 12.0 Hz, J_2 = 3.4 Hz, 1H, $H_{5'}$), 4.25 (dd, J_1 = 12.1 Hz, J_2 = 5.1 Hz, 1H, $H_{5'}$), 2.88-2.81 (m, 1H, $H_{2'}$), 2.66-2.60 (m, 1H, $H_{2'}$), 2.14-2.11 (m, 6H); **24β**: δ: 9.69 (s, br, 1H), 8.31 (d, J_1 = 2.5 Hz, 1H), 7.70 (dd, J_1 = 9.0 Hz, J_2 = 2.5 Hz, 1H), 7.50 (d, J_1 = 9.0 Hz, 1H), 6.68 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.41-5.37 (m, 1H, $H_{3'}$), 4.46-4.38 (m, 2H, $H_{5'}$ and $H_{5'}$), 4.20 (q, J_1 = 3.8 Hz, 1H, $H_{4'}$), 3.03 (dt, J_1 = 16.4 Hz, J_2 = 8.3 Hz, 1H, $H_{2'}$), 2.24-2.19 (m, 1H, $H_{2'}$), 2.14-2.11 (m, 6H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 170.7, 170.65, 170.63, 170.5, 160.8, 160.7, 150.0, 149.8, 138.8, 138.4, 137.77, 137.75, 131.55, 131.53, 118.5, 118.0, 117.9, 117.2, 117.1, 85.3, 84.6, 81.4, 81.1, 74.1, 73.3, 64.2, 63.6, 34.4, 34.3, 21.07, 21.06, 21.04, 21.02; **IR** (neat) ν: 3005, 1733, 1715, 1699, 1220, 1050, 772 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{17}H_{17}BrN_2O_7Na$, 463.01168; found 463.01108.

1'-(3'-*O*-Acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-α-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (25α) and **1'-(3'-*O*-acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (25β)**



25α **25β**

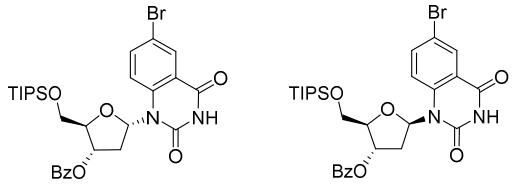
According to the general procedure A: The separable isomer mixture of nucleosides **25α** (0.766 g, 30%) and **25β** (1.352 g, 54%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 8:2 → 6:4) using 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21** (1.312 g, 5.44 mmoles), BSA (2.77 mL, 11.34 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 1.99 g, 4.54 mmoles), NIS (1.22 g, 5.44 mmoles) and TMSOTf (0.49 mL, 2.72 mmoles). The reaction was also performed on a large scale using 6-bromo-quinazoline-2,4-(1*H*,3*H*)-dione **21** (6.593 g, 27.354 mmoles), BSA (13.93 mL, 56.988 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 10.00 g, 22.795 mmoles), NIS (6.154 g, 27.354 mmoles) and TMSOTf (2.47 mL, 13.677 mmoles) to give nucleosides **25α** (3.620 g, 29%) and **25β** (6.520 g, 52%).

Alternative procedure: (Starting from imidate **47**): To a suspension of commercially available 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.078 g, 0.325 mmoles) in CH_2Cl_2 (0.06 M) with activated molecular sieves (MS 4°A, twice the weight of the donor) was added BSA (0.17 mL, 0.677 mmoles) dropwise over a 5 min period. The solution was stirred at room temperature for 1h, (the complete dissolution was observed after 0.5h). To this solution was added trichloroacetimidate **47** at 0°C. After 5 min, TMSOTf (0.059 mL, 0.325 mmoles) was added and the reaction mixture was stirred for 10min at 0°C. It was then quenched with aq. sat. sol. $NaHCO_3$ and extracted with CH_2Cl_2 (3 times). The combined organics layers were dried over $MgSO_4$, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane:EtOAc, 7:3 then 6:4) to afford the separable isomer mixture of nucleosides **25α** (0.037 g, 25%) and **25β** (0.067 g, 44%) as off-white foams.

Characterization: R_f (Hexane/EtOAc, 7:3): **25α**: 0.28; **25β**: 0.14; **^{1}H NMR** (2D COSY and ROESY, 500 MHz, $CDCl_3$) **25α**: δ: 9.75 (s, br, 1H), 8.32 (d, J_1 = 1.6 Hz, 1H), 7.68 (dd, J_1 = 9.1 Hz, J_2 = 1.6 Hz, 1H), 7.64 (d, J_1 = 9.1 Hz, 1H), 7.05 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.52 (d, J_1 = 8.5 Hz, 1H, $H_{3'}$), 4.38 (s,

1H, $H_{4'}$), 4.00 (d, J_1 = 9.7 Hz, 1H, $H_{5'}$), 3.94 (dd, J_1 = 10.8 Hz, J_2 = 2.1 Hz, 1H, $H_{5'}$), 2.87-2.81 (m, 1H, $H_{2'}$), 2.43 (ddd, J_1 = 14.6 Hz, J_2 = 7.6 Hz, J_3 = 3.8 Hz, 1H, $H_{2'}$), 2.14 (s, 1H), 1.11-1.08 (m, 21H); **25β**: δ: 9.39 (s, br, 1H), 8.32 (d, J_1 = 2.5 Hz, 1H), 7.87 (d, J_1 = 9.1 Hz, 1H), 7.62 (dd, J_1 = 9.1 Hz, J_2 = 2.5 Hz, 1H), 6.82 (dd, J_1 = 9.5 Hz, J_2 = 6.2 Hz, 1H, $H_{1'}$), 5.52-5.50 (m, 1H, $H_{3'}$), 4.11-4.03 (m, 3H, $H_{4'}$, $H_{5'}$ and $H_{5'}$), 2.85-2.81 (m, 1H, $H_{2'}$), 2.16-2.14 (m, 1H, $H_{2'}$), 2.14 (s, 3H), 1.21-1.07 (m, 21H); **^{13}C NMR** (125 MHz, $CDCl_3$) **25α**: δ: 170.7, 160.9, 150.0, 138.1, 137.3, 131.3, 118.6, 116.8, 86.2, 84.8, 75.7, 65.7, 35.2, 21.2, 18.13, 18.11, 12.0; **25β**: δ: 170.8, 160.8, 150.3, 138.3, 137.8, 131.2, 119.4, 118.6, 117.17 84.2, 84.0, 72.7, 62.6, 33.9, 21.1, 18.14, 18.13, 12.0; **IR** (neat) **25α**: ν: 3191, 3072, 2941, 2865, 2360, 2341, 1741, 1706, 1691, 1603, 1483, 1462, 1360, 1311, 1229, 1015, 882, 681, 503; **25β**: ν: 3196, 3064, 2942, 2865, 2360, 2341, 1739, 1698, 1603, 1483, 1465, 1361, 1318, 1236, 1015, 881, 680, 504 cm^{-1}; **HR-ESI MS** (*m/z*): $[M+Na]^+$ calcd for $C_{24}H_{35}BrN_2O_6SiNa$, 577.13455; found **25α**: 577.13345; **25β**: 577.13335.

1'-(3'-*O*-Benzoyl-5'-*O*-triisopropylsilyl-2'-deoxy-α-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (26α) and **1'-(3'-*O*-benzoyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (26β)**



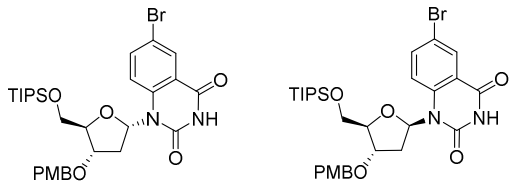
26α 26β

According to the general procedure A: The separable isomer mixture of nucleosides **26α** (0.075 g, 32%) and **26β** (0.120 g, 51%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 8:2 → 7:3) using 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.110 g, 0.455 mmoles), BSA (0.23 mL, 0.948 mmoles), thioglycoside **15** (α/β = 1.0:1.8, 0.190 g, 0.379 mmoles), NIS (0.102 g, 0.455 mmoles) and TMSOTf (41 μL, 0.228 mmoles).

Characterization: R_f (Hexane/EtOAc, 7:3): **26α**: 0.40; **26β**: 0.27; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **26α**: δ: 9.35 (s, br, 1H), 8.34 (d, J_1 = 2.3 Hz, 1H), 8.08 (d, J_1 = 7.8 Hz, 2H), 7.78 (d, J_1 = 9.1 Hz, 1H), 7.64 (t, J_1 = 7.5 Hz, 1H), 7.59 (dd, J_1 = 9.1 Hz, J_2 = 2.4 Hz, 1H), 7.52 (t, J_1 = 7.7 Hz, 2H), 7.19 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.77 (d, J_1 = 7.9 Hz, 1H, $H_{3'}$), 4.56 (s, 1H, $H_{4'}$), 4.15 (d, J_1 = 9.3 Hz, 1H, $H_{5'}$), 4.03 (dd, J_1 = 10.6 Hz, J_2 = 2.0 Hz, 1H, $H_{5'}$), 3.03-2.95 (m, 1H, $H_{2'}$), 2.62 (ddd, J_1 = 15.8 Hz, J_2 = 7.3 Hz, J_3 = 3.2 Hz, 1H, $H_{2'}$), 1.15-1.13 (m, 21H); **26β**: δ: 9.79 (s, br, 1H), 8.33 (d, J_1 = 2.1 Hz, 1H), 8.06 (d, J_1 = 8.0 Hz, 2H), 7.91

(d, J_1 = 9.1 Hz, 1H), 7.65 (dd, J_1 = 8.7 Hz, J_2 = 2.0 Hz, 1H), 7.58 (t, J_1 = 7.6 Hz, 1H), 7.45 (t, J_1 = 7.6 Hz, 2H), 6.92 (dd, J_1 = 9.3 Hz, J_2 = 6.3 Hz, 1H, $H_{1'}$), 5.79-5.76 (m, 1H, $H_{3'}$), 4.22-4.11 (m, 3H, $H_{4'}$, $H_{5'}$ and $H_{5'}$), 3.02-2.94 (m, 1H, $H_{2'}$), 2.30 (dd, J_1 = 15.3 Hz, J_2 = 6.6 Hz, 1H, $H_{2'}$), 1.16-1.11 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) **26α**: δ: 166.3, 160.7, 150.0, 138.1, 137.5, 133.8, 131.4, 129.7, 129.6, 128.8, 118.8, 118.7, 116.9, 86.4, 85.2, 76.6, 66.1, 35.4, 18.2, 18.1, 12.0; **26β**: δ: 166.2, 160.8, 150.2, 138.5, 137.8, 133.5, 131.2, 129.9, 129.6, 128.6, 119.4, 118.6, 117.1, 84.3, 84.2, 73.3, 62.8, 34.2, 18.19, 18.17, 12.1; **IR** (neat) **26α**: ν: 3186, 3071, 2942, 2865, 1716, 1692, 1603, 1483, 1463, 1314, 1269, 1248, 1095, 773, 710; **26β**: ν: 3187, 3066, 2943, 2865, 1713, 1603, 1485, 1467, 1316, 1270, 1105, 773, 711 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{29}H_{37}BrN_2O_6SiNa$, 639.15020; found **26α**: 639.14905; **26β**: 639.14948.

1'-(3'-*O*-*p*-Methoxybenzoyl-5'-*O*-triisopropylsilyl-2'-deoxy-α-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (27α) and **1'-(3'-*O*-*p*-methoxybenzoyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (27β)**

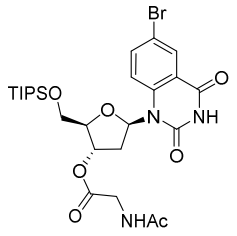


According to the general procedure A: The separable isomer mixture of nucleosides **27α** (0.072 g, 30%) and **27β** (0.123 g, 50%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 8:2 → 6:4) using 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.109 g, 0.452 mmoles), BSA (0.23 mL, 0.942 mmoles), thioglycoside **16** (α/β = 1.0:1.8, 0.20 g, 0.377 mmoles), NIS (0.102 g, 0.452 mmoles) and TMSOTf (62 μL, 0.226 mmoles).

Characterization: 27α: R_f (Hexane/EtOAc, 7:3): 0.38; **^{1}H-NMR** (2D COSY, 400 MHz, $CDCl_3$) δ: 8.69 (s, br, 1H), 8.33 (d, J_1 = 2.3 Hz, 1H), 8.02 (d, J_1 = 8.8 Hz, 2H), 7.78 (d, J_1 = 9.1 Hz, 1H), 7.61 (dd, J_1 = 9.1 Hz, J_2 = 2.4 Hz, 1H), 7.17 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 6.99 (d, J_1 = 8.8 Hz, 2H), 5.73 (d, J_1 = 7.9 Hz, 1H, $H_{3'}$), 4.54 (s, 1H, $H_{4'}$), 4.14 (dd, J_1 = 10.8 Hz, J_2 = 1.8 Hz, 1H, $H_{5'}$), 4.03 (dd, J_1 = 10.8 Hz, J_2 = 2.1 Hz, 1H, $H_{5'}$), 3.91 (s, 3H), 3.00-2.92 (m, 1H, $H_{2'}$), 2.59 (ddd, J_1 = 14.9 Hz, J_2 = 7.5 Hz, J_3 = 3.2 Hz, 1H, $H_{2'}$), 1.15-1.12 (m, 21H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 166.1, 164.1, 160.4, 149.8, 138.1, 137.6, 131.8, 131.4, 121.9, 118.9, 118.7, 116.9, 114.1, 86.5, 85.3, 76.3, 66.2, 55.7, 35.5, 18.21, 18.19, 12.0; **IR** (neat) ν: 2960, 2940, 2866, 1715, 1604, 1462, 1258, 1095, 772, 418 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for

$C_{30}H_{39}BrN_2O_7SiNa$, 669.16076; found 669.15957; **27β**: R_f (Hexane/EtOAc, 7:3): 0.23; **^{1}H-NMR** (2D COSY, 400 MHz, $CDCl_3$) δ: 8.94 (s, br, 1H), 8.33 (d, J_1 = 2.3 Hz, 1H), 8.01 (d, J_1 = 8.8 Hz, 2H), 7.92 (d, J_1 = 9.1 Hz, 1H), 7.65 (dd, J_1 = 9.0 Hz, J_2 = 2.3 Hz, 1H), 6.93 (d, J_1 = 8.8 Hz, 1H), 6.89 (dd, J_1 = 9.8 Hz, J_2 = 6.5 Hz, 1H, $H_{1'}$), 5.75-5.72 (m, 1H, $H_{3'}$), 4.19-4.12 (m, 3H, $H_{4'}$, $H_{5'}$ and $H_{5'}$), 3.88 (s, 3H), 2.99-2.91 (m, 1H, $H_{2'}$), 2.27 (dd, J_1 = 14.1 Hz, J_2 = 6.4 Hz, 1H, $H_{2'}$), 1.20-1.11 (m, 21H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 165.9, 163.9, 160.5, 149.9, 138.5, 137.9, 132.0, 131.3, 122.0, 119.5, 118.6, 117.2, 113.9, 84.5, 84.3, 73.0, 62.9, 55.6, 34.3, 18.23, 18.21, 12.1; **IR** (neat) ν: 2960, 2940, 2866, 1715, 1605, 1257, 1218, 1065, 772, 417 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{30}H_{39}BrN_2O_7SiNa$, 669.16076; found 669.15968.

1'-(3'-*O*-(*N*-Acetyl)-glycyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (28β)

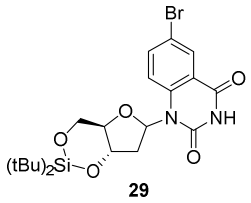


28β

According to the general procedure A: Nucleoside **28β** (0.150 g, 41%) was obtained as an off-white foam after column chromatography on silica gel (Hexane/EtOAc, 1:9 → 0:1) using 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.175 g, 0.726 mmoles), BSA (0.37 mL, 1.513 mmoles), enantiopure thioglycoside **17α** or **17β** (0.30 g, 0.605 mmoles), NIS (0.164 g, 0.726 mmoles) and TMSOTf (66 μL, 0.363 mmoles).

Characterization: R_f (Hexane/EtOAc, 1:9): 0.24; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) δ: 10.06 (s, 1H), 8.19 (d, J_1 = 2.5 Hz, 1H), 7.80 (d, J_1 = 9.1 Hz, 1H), 7.59 (dd, J_1 = 9.1 Hz, J_2 = 2.5 Hz, 1H), 6.73 (dd, J_1 = 9.3 Hz, J_2 = 6.2 Hz, 1H, $H_{1'}$), 6.45 (t, J_1 = 5.3 Hz, 1H), 5.59-5.56 (m, 1H, $H_{3'}$), 4.08 (d, J_1 = 5.4 Hz, 2H), 4.06-4.01 (m, 3H, $H_{4'}$, $H_{5'}$ and $H_{5'}$), 2.90-2.82 (m, 1H, $H_{2'}$), 2.19 (ddd, J_1 = 14.1 Hz, J_2 = 6.2 Hz, J_3 = 2.1 Hz, 1H, $H_{2'}$), 2.05 (s, 3H), 1.16-1.07 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 170.7, 169.9, 160.8, 150.2, 138.5, 137.7, 131.0, 119.2, 118.4, 117.0, 84.1, 84.0, 74.0, 62.7, 41.6, 33.9, 22.9, 18.11, 18.09, 12.0; **IR** (neat) ν: 3225, 2943, 2866, 1706, 1603, 1484, 1467, 1187 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{26}H_{38}BrN_3O_7SiNa$, 634.15601; found 634.15505.

1'-(3',5'-*O*-(Di-*tert*-butylsilylene)-2'-deoxy-α,β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (29α) and **(29β)**

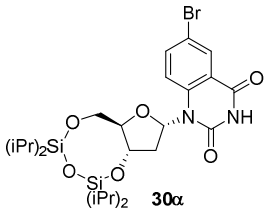


29

According to the general procedure A: The inseparable isomer mixture of nucleosides **29α** and **29β** (α/β = 1.8:1.0, 0.164 g, 50%) were obtained as a white solid after column chromatography on silica gel (Hexane/EtOAc, 8:2) using 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.190 g, 0.788 mmoles), BSA (0.40 mL, 1.642 mmoles), thioglycoside **18** (α/β = 1.0:1.8, 0.350 g, 0.657 mmoles), NIS (0.177 g, 0.788 mmoles) and TMSOTf (71 μL, 0.394 mmoles).

Characterization: $\mathbf{R}_f$ (Hexane/EtOAc, 8:2): 0.23; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **29α**: δ: 9.77 (s, br, 1H), 8.35 (d, J_1 = 2.5 Hz, 1H), 7.75 (d, J_1 = 2.5 Hz, 1H), 7.29 (d, J_1 = 9.2 Hz, 1H), 6.45 (dd, J_1 = 9.2 Hz, J_2 = 6.4 Hz, 1H, $H_{1'}$), 4.50-4.33 (m, 3H, $H_{3'}$, $H_{4'}$ and $H_{5'}$), 3.92 (dd, J_1 = 9.7 Hz, J_2 = 8.9 Hz, 1H, $H_{5'}$), 3.09-3.02 (m, 1H, $H_{2'}$), 2.53 (dt, J_1 = 12.8 Hz, J_2 = 6.5 Hz, 1H, $H_{2'}$), 1.13-1.06 (m, 18H); **29β**: δ: 9.80 (s, br, 1H), 8.33 (d, J_1 = 2.5 Hz, 1H), 7.73 (d, J_1 = 2.5 Hz, 1H), 7.37 (d, J_1 = 9.1 Hz, 1H), 6.39 (dd, J_1 = 9.6 Hz, J_2 = 3.6 Hz, 1H, $H_{1'}$), 4.81 (q, J_1 = 8.8 Hz, 1H, $H_{3'}$), 4.50-3.37 (m, 1H, $H_{5'}$), 4.10 (t, J_1 = 10.4 Hz, 1H, $H_{5'}$), 3.71 (ddd, J_1 = 10.4 Hz, J_2 = 8.9 Hz, J_3 = 5.0 Hz, 1H, $H_{4'}$), 2.90 (ddd, J_1 = 6.5 Hz, J_2 = 3.5 Hz, J_3 = 2.1 Hz, 1H, $H_{2'}$), 2.32-2.28 (m, 1H, $H_{2'}$), 1.13-1.06 (m, 18H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 161.1, 161.0, 149.6, 149.1, 139.9, 139.7, 138.18, 138.15, 131.7, 131.5, 118.3, 118.2, 117.1, 116.9, 116.3, 116.5, 84.9, 84.2, 78.4, 77.6, 76.2, 75.5, 68.2, 67.5, 35.8, 34.4, 27.7, 27.6, 27.4, 27.3, 22.85, 22.82, 20.4, 20.3; **IR** (neat) ν: 2967, 2934, 2893, 2859, 1700, 1603, 1472, 1316, 1052, 826, 772, 749, 426 cm^{-1}; **HR-ESI MS** (m/z): $[M+Na]^+$ calcd for $C_{21}H_{29}BrN_2O_5SiNa$, 519.09268; found 519.09174.

1'-(3',5'-*O*-(Tetraisopropylsiloxane-1,3-diyl)-2'-deoxy-α-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3H)-dione (30α)

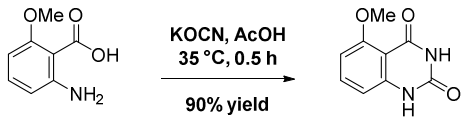


30α

According to the general procedure A: Nucleosides **30α** (0.139 g, 56%) was obtained as an off-white foam after column chromatography on silica gel (Hexane/EtOAc, 9:1 → 8:2) using 6-bromo-quinazoline-2,4-(1H,3H)-dione **21** (0.120 g, 0.497 mmoles), BSA (0.26 mL, 1.035 mmoles), thioglycoside **19** (α/β = 1.0:1.8, 0.200 g, 0.414 mmoles), NIS (0.112 g, 0.497 mmoles) and TMSOTf (45 μL, 0.248 mmoles).

Characterization: R_f (Hexane/EtOAc, 8:2): 0.23; **^{1}H NMR** (2D COSY and ROESY, 500 MHz, $CDCl_3$) δ: 9.53 (s, br, 1H), 8.34 (d, J_1 = 2.5 Hz, 1H), 7.75 (dd, J_1 = 9.1 Hz, J_2 = 2.5 Hz, 1H), 7.49 (d, J_1 = 9.1 Hz, 1H), 6.68 (dd, J_1 = 9.8 Hz, J_2 = 6.6 Hz, 1H, $H_{1'}$), 4.69 (q, J_1 = 7.5 Hz, 1H, $H_{3'}$), 4.29 (sext, J_1 = 3.6 Hz, 1H, $H_{4'}$), 4.02 (dd, J_1 = 12.1 Hz, J_2 = 3.6 Hz, 1H, $H_{5'}$), 3.88 (dd, J_1 = 12.1 Hz, J_2 = 7.0 Hz, 1H, $H_{5'}$), 2.86-2.79 (m, 1H, $H_{2'}$), 2.57-2.52 (m, 1H, $H_{2'}$), 1.11-1.04 (m, 28H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 160.8, 150.1, 138.5, 137.9, 131.5, 118.6, 117.9, 117.0, 84.3, 84.1, 72.9, 63.5, 36.3, 17.7, 17.62, 17.60, 17.5, 17.4, 17.3, 17.2, 17.1, 13.6, 13.4, 13.1, 12.7; **IR** (neat) ν: 2943, 2867, 1708, 1603, 1483, 1465, 1315, 1143, 1118, 1091, 885, 774, 705 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{25}H_{39}BrN_2O_6Si_2Na$, 621.14277; found 621.14273.

5-Methoxyquinazolin-2,4-(1*H*,3*H*)-dione (32)



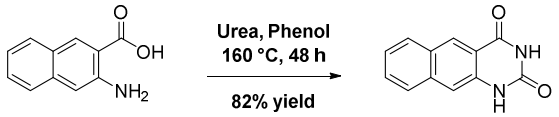
41 32

Procedure: Following the literature procedure,[114] 2-amino-6-methoxybenzoic acid **41** (0.85 g, 5.09 mmoles) was suspended in aqueous AcOH (32 mL, 98%). The slurry was stirred at 35 °C for 15 min. A solution of KOCN (1.03 g, 12.71 mmoles) in H_2O (3.4 mL) was added dropwise to the reaction mixture. The mixture for stirred at 35 °C for 0.5h. NaOH (9.1 g) was added slowly to the reaction which was then cooled to room temperature. The pH was adjusted to 4.0 with conc. aq. HCl. The precipitate was filtered, washed with H_2O and the white solid was dried at 60 °C for 18h to provide **32** as a white solid (0.88 g, 90%).

Characterization: **^{1}H NMR** (500 MHz, DMSO-d_6) δ: 10.96 (s, br, 1H), 10.86 (s, br, 1H), 7.49 (s, 1H), 6.70 (s, 2H), 3.81 (s, 3H); **^{13}C-NMR** (125 MHz, DMSO-d_6) δ: 160.8, 160.4, 150.1, 143.1, 135.2, 107.3, 104.9, 103.7, 55.8; **IR** (neat) ν: 3150, 3052, 3011, 2835, 1732, 1675, 1614, 1595, 1515, 1431, 1408,

1260, 1230, 1112, 984, 802, 766, 751 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_9H_8N_2O_3Na$ 215.04326, found 215.04284.

Benzo[g]quinazoline-2,4-(1*H*,3*H*)-dione (33)

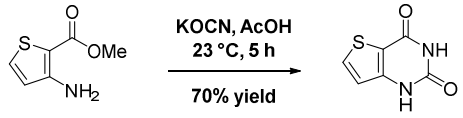


42 33

Procedure: Following the literature procedure,[156] 3-amino-2-naphtoic acid **42** (2.00 g, 10.68 mmoles) and urea (3.98 g, 66.24 mmoles) were heated at 160˚C for 48h in phenol (35.19 g, 373.94 mmoles). The solution was then cooled to 80 ˚C and poured into H_2O (100 mL). The solution was filtered and washed with H_2O. The product was dissolved in DMF (40 mL) and heated at 100 ˚C for 0.5h. $CHCl_3$ (100 mL) was then added and the mixture was cooled overnight at 4 ˚C. The solid was collected by filtration and dried under high vacuum to give **33** as an ochre powder (1.85 g, 82%).

Characterization: R_f (MeOH/$CHCl_3$, 1:9): 0.42; **^{1}H NMR** (500 MHz, DMSO-d_6) δ: 11.22 (s, br, 2H), 8.62 (s, 1H), 8.06 (d, J_1 = 8.0 Hz, 1H), 7.85 (d, J_1 = 8.0 Hz, 1H), 7.57 (td, J_1 = 8.0 Hz, J_2 = 1.2 Hz, 1H), 7.51 (s, 1H), 7.42 (td, J_1 = 8.0 Hz, J_2 = 1.0 Hz, 1H); **^{13}C NMR** (125 MHz, DMSO-d_6) δ: 162.8, 150.3, 136.5, 136.3, 129.5, 129.2, 128.9, 128.4, 126.7, 124.9, 115.1, 110.2; **IR** (neat) *v*: 3345, 3213, 3151, 3025, 2830, 1699, 1665, 1633, 1580, 1524, 1474, 1433, 1319, 1296, 1215, 872, 846, 778, 740 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + H]^+$ calcd for $C_{12}H_9N_2O_2$ 213.06640, found 213.06569.

Thieno[3,2-*d*]pyrimidine-2,4-(1*H*,3*H*)-dione (34)



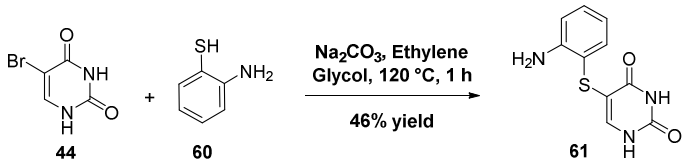
43 34

Procedure: Following the literature procedure,[142] a solution of KOCN (2.06 g, 25.45 mmoles), in H_2O (4.5 mL) was added dropwise to a mixture of methyl 3-aminothiophene-2-carboxylate **43** (2.00 g, 12.72 mmoles), in aqueous AcOH (32 mL, 50%). The reaction mixture for stirred for 5h at room temperature. The precipitate which has formed was collected by filtration and then dissolved in 2N NaOH solution (32 mL). The solution was stirred at room temperature for 1h and the slurry was filtered to remove any undissolved material. The solution was then acidified with conc. aq. HCl until a pH 5-6 was

obtained. The precipitate was filtered and the solid was dried at 50 °C for 18h to provide **34** as a white solid (1.49 g, 70%).

Characterization: **^{1}H NMR** (500 MHz, DMSO-d_6) δ: 8.04 (s, 1H), 6.93 (s, 1H); **^{13}C NMR** (125 MHz, DMSO-d_6) δ: 159.2, 151.7, 146.8, 135.9, 117.4, 111.2; **IR** (neat) ν: 3466, 3379, 3092, 2865, 2801, 1665, 1645, 1574, 1538, 1443, 1365, 1286, 1119, 844, 778, 749 cm^{-1}; **HR-ESI MS** (*m*/*z*): [M + H]$^+$ calcd for $C_6H_5N_2O_2S$ 169.00718, found 169.00669.

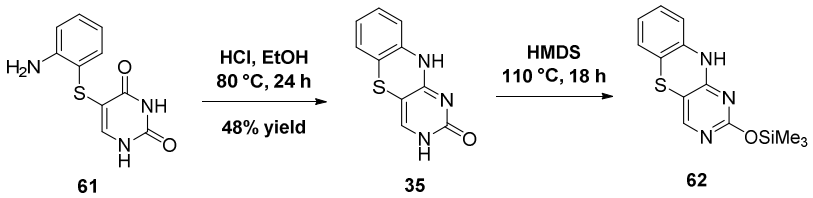
5-(2-aminophenylthio)pyrimidine-2,4-(1*H*,3*H*)-dione (61)



Procedure: Following the literature procedure,[309] 2-aminothio-phenol **60** (3.51 g, 28.07 mmoles), 5-bromouracil **44** (4.00 g, 20.94 mmoles) and Na_2CO_3 (2.22 g, 20.94 mmoles) were suspended in ethylene glycol (12.7 mL). The reaction mixture was stirred for 1h at 120 °C and then cooled to room temperature. A precipitate was formed. While stirring, H_2O (42 mL) was added and the solution was neutralized by the addition of AcOH (1.4 mL). The solid was filtered off, washed with H_2O and a large volume of EtOH until the initially dark yellow solid became colorless. It was then washed with Et_2O to facilitate drying. A solution of NaOH (0.92 g) in H_2O (46 mL) was heated at 50 °C and the previous white solid was added in portions. When everything was dissolved, AcOH (1.7 mL, until pH = 7) was added to neutralize the clear solution. A white precipitate was formed. The hot suspension was filtered off, washed with H_2O, EtOH and Et_2O to give **61** as a white solid (2.25 g, 46%).

Characterization: **^{1}H NMR** (500 MHz, DMSO-d_6) δ: 11.37 (s, 1H), 11.15 (s, 1H), 7.52 (s, 1H), 7.29 (dd, J_1 = 7.7 Hz, J_2 = 1.4 Hz, 1H), 7.04 (td, J_1 = 8.1 Hz, J_2 = 1.5 Hz, 1H), 6.68 (dd, J_1 = 8.1 Hz, J_2 = 1.0 Hz, 1H), 6.49 (td, J_1 = 7.6 Hz, J_2 = 1.2 Hz, 1H), 5.51 (s, 1H); **^{13}C NMR** (125 MHz, DMSO-d_6) δ: 163.0, 151.0, 149.6, 144.3, 135.4, 130.0, 116.4, 114.7, 114.5, 105.7; **IR** (neat) ν: 3431, 3351, 3221, 3002, 2815, 1707, 1666, 1607, 1478, 1429, 1301, 1234, 1149, 845, 746 cm^{-1}; **ESI MS** (*m*/*z*): [M + H]$^+$ calcd for $C_{10}H_{10}N_3O_2S$ 236.0, found 235.9.

1,3-Diaza-2-oxophenothiazine (35)

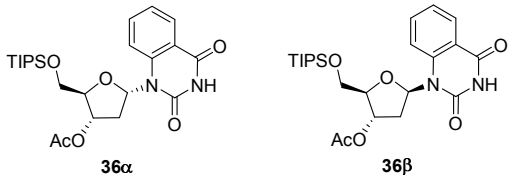


61 35 62

Procedure: Following the literature procedure,[309] HCl (3.73 mL, 37%) was added to a flask containing EtOH (38 mL). To this solution, was added **61** (1.80 g, 7.65 mmoles) in portions and the reaction mixture was stirred at 80 °C for 24h. A white precipitate of **61**·HCl formed first after **61** had dissolved. The precipitate gradually changed color as the product **35**·HCl was formed as a mass of thin yellow crystals. The reaction mixture was cooled to room temperature and filtered. To prepare the free base of **35** (**35**·HCl → **35**), the previous precipitate was added to a 50 °C aqueous 5% ammonia solution (40 mL). The yellow crystals turned to a pale yellow precipitate. After 5 min, the hot mixture was filtered and the recovered solid was added again to a 50 °C aqueous 5% ammonia solution (40 mL). After 5 min, it was filtered, washed with H_2O, Et_2O and dried at 60 °C under reduce pressure to give **35** as a yellow solid (0.79 g, 48%). To obtain ^1H and ^{13}C NMR data, **35** (0.119 g, 0.547 mmoles) was suspended in HMDS (4 mL) and stirred at 110°C for 18h. The solvent was evaporated and the residue was dried reduce pressure to give the corresponding silylated compound **62** as a yellow solid.

Characterization: **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 7.67 (s, 1H), 7.14 (s, br, 1H), 6.99 (t, J_1 = 7.3 Hz, 1H), 6.87 (m, 2H), 6.58 (d, J_1 = 7.6 Hz, 1H), 0.34 (s, 9H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.4, 160.4, 152.6, 137.8, 127.7, 126.9, 124.5, 117.4, 116.1, 104.2, 0.4; **IR** (neat) ν: 3238, 1557, 1408, 1249, 1019, 848, 757 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + H]^+$ calcd for $C_{10}H_8N_3OS$ 218.03880, found 218.03824.

1'-(3'-*O*-Acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-α-D-ribofuranoside)-quinazoline-2,4-(3H)-dione (36α) and **1'-(3'-*O*-acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-quinazoline-2,4-(3H)-dione (36β)**

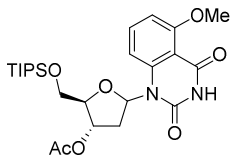


36α 36β

According to the general procedure A: The separable isomer mixture of nucleosides **36α** (0.065 g, 30%) and **36β** (0.116 g, 53%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 8:2 → 6:4) using quinazoline-2,4-(1H,3H)-dione **31** (0.089 g, 0.547 mmoles), BSA (0.28 mL, 1.14 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 0.20 g, 0.456 mmoles), NIS (0.113 g, 0.502 mmoles) and HOTf (16 μL, 0.182 mmoles).

Characterization: R_f (Hexane/EtOAc, 7:3): **36α**: 0.29; **36β**: 0.22; **^{1}H NMR** (2D COSY and ROESY, 400 MHz, $CDCl_3$) **36α**: δ: 9.17 (s, br, 1H), 8.22 (dd, J_1 = 7.9 Hz, J_2 = 1.6 Hz, 1H), 7.73 (d, J_1 = 7.8 Hz, 1H), 7.64-7.60 (m, 1H), 7.29-7.25 (m, 1H), 7.05 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.53 (ddd, J_1 = 8.5 Hz, J_2 = 4.0 Hz, J_3 = 2.5 Hz, 1H, $H_{3'}$), 4.42 (d, J_1 = 2.4 Hz, 1H, $H_{4'}$), 4.01 (dd, J_1 = 10.8 Hz, J_2 = 2.5 Hz, 1H, $H_{5'}$), 3.96 (dd, J_1 = 10.8 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 2.87-2.80 (m, 1H, $H_{2'}$), 2.54 (ddd, J_1 = 14.5 Hz, J_2 = 7.6 Hz, J_3 = 4.0 Hz, 1H, $H_{2'}$), 2.14 (s, 3H), 1.16-1.08 (m, 21H); **36β**: δ: 9.54 (s, br, 1H), 8.22 (dd, J_1 = 7.9 Hz, J_2 = 1.6 Hz, 1H), 7.95 (d, J_1 = 8.6 Hz, 1H), 7.60-7.56 (m, 1H), 7.28-7.25 (m, 1H), 6.85 (dd, J_1 = 9.5 Hz, J_2 = 6.7 Hz, 1H, $H_{1'}$), 5.54 (ddd, J_1 = 7.9 Hz, J_2 = 4.1 Hz, J_3 = 2.2 Hz, 1H, $H_{3'}$), 4.09-4.03 (m, 3H, $H_{4'}$, $H_{5'}$ and $H_{5'}$), 2.97-2.89 (m, 1H, $H_{2'}$), 2.14 (dd, J_1 = 6.3 Hz, J_2 = 2.2 Hz, 1H, $H_{2'}$), 2.10 (s, 3H), 1.12-1.09 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) **36α**: δ: 170.8, 161.9, 150.2, 139.3, 134.6, 129.0, 123.7, 117.0, 116.7, 86.1, 84.6, 75.7, 65.6, 35.1, 21.2, 18.16, 18.14, 12.0; **36β**: δ: 170.8, 161.9, 150.6, 139.5, 135.0, 128.8, 123.9, 117.5, 116.9, 84.2, 84.0, 72.9, 62.7, 33.9, 21.1, 18.17, 18.15, 12.1; **IR** (neat) **36α**: ν: 3210, 3065, 2942, 2856, 2363, 2363, 1704, 1687, 1609, 1483, 1386, 1313, 1231, 772, 686; **36β**: ν: 3210, 3066, 2942, 2865, 1740, 1690, 1609, 1483, 1386, 1314, 1239, 772, 686 cm^{-1}; **HR-ESI MS** (m/z): $[M+Na]^+$ calcd for $C_{24}H_{36}N_2O_6SiNa$, 499.22403; **36α**: found 499.22307; **36β**: found 499.22339.

1'-(3'-*O*-Acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-α,β-D-ribofuranoside)-5-methoxy-quinazoline-2,4-(3H)-dione (37β) and **(37α)**



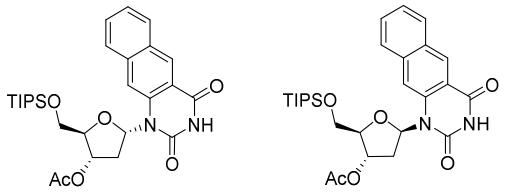
37

According to the general procedure A: The inseparable isomer mixture of nucleosides **37α** and **37β** (α/β = 1.0:1.7, 0.184 g, 80%) were obtained as an off-white foam after column chromatography on silica gel (Hexane/EtOAc, 4:6 → 3:7) using 5-methoxyquinazoline-2,4-(3H)-dione **32** (0.105 g, 0.547

mmoles), BSA (0.28 mL, 1.14 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 0.20 g, 0.456 mmoles), NIS (0.112 g, 0.502 mmoles) and HOTf (16 μL, 0.182 mmoles).

Characterization: R_f (Hexane/EtOAc, 3:7): 0.25; **^{1}H NMR** (2D COSY and ROESY, 500 MHz, $CDCl_3$) **37α:** δ: 9.24 (s, 1H), 7.46 (t, J_1 = 8.6 Hz, 1H), 7.21 (d, J_1 = 8.7 Hz, 1H), 6.89 (t, J_1 = 7.9 Hz, 1H, $H_{1'}$), 6.74-6.70 (m, 1H), 5.47-5.45 (m, 1H, $H_{3'}$), 4.37-4.36 (m, 1H, $H_{4'}$), 4.00-3.90 (m, 2H, $H_{5'}$ and $H_{5'}$), 3.89 (s, 3H), 2.77-2.10 (m, 1H, $H_{2'}$), 2.54 (ddd, J_1 = 14.4 Hz, J_2 = 7.7 Hz, J_3 = 4.3 Hz, 1H, $H_{2'}$), 2.08 (s, 3H), 1.13-1.05 (s, 21H); **37β**: δ: 9.36 (s, 1H), 7.44-7.42 (m, 2H), 6.74-6.70 (m, 2H, $H_{1'}$), 5.47-5.45 (m, 1H, $H_{3'}$), 4.00-3.90 (m, 3H, $H_{4'}$, $H_{5'}$ and $H_{5'}$), 3.89 (s, 3H), 2.95-2.89 (m, 1H, $H_{2'}$), 2.04 (s, 3H), 2.04-1.99 (m, 1H, $H_{2'}$), 1.13-1.05 (m, 21H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 178.6, 170.77, 170.75, 170.6, 161.6, 161.5, 160.4, 160.3, 150.3, 150.15, 150.14, 141.9, 141.7, 135.15, 135.12, 134.8, 109.4, 108.7, 86.3, 84.25, 84.24, 83.8, 75.3, 72.9, 65.2, 62.6, 56.4, 56.3, 34.7, 33.6, 21.0, 20.9, 17.99, 17.98, 11.9, 11.8; **IR** (neat) ν: 3221, 2942, 2865, 1706, 1599, 1489, 1270, 772 cm^{-1}; **HR-ESI MS** (*m/z*): $[M+Na]^+$ calcd for $C_{25}H_{38}N_2O_7SiNa$, 529.23460; found 529.23382.

1'-(3'-*O*-Acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-α-D-ribofuranoside)-benzo[g]quinazoline-2,4-(3H)-dione (38α) and **1'-(3'-*O*-acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-benzo[g]quinazoline-2,4-(3H)-dione (38β)**



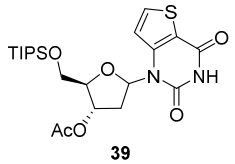
38α **38β**

According to the general procedure A: The separable isomer mixture of nucleosides **38α** (0.060 g, 25%) and **38β** (0.107 g, 45%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 8:2 → 7:3) using benzo[g]quinazoline-2,4-(3H)-dione **33** (0.116 g, 0.547 mmoles), BSA (0.28 mL, 1.14 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 0.20 g, 0.456 mmoles), NIS (0.113 g, 0.502 mmoles) and HOTf (16 μL, 0.182 mmoles).

Characterization: R_f (Hexane/EtOAc, 7:3): **38α**: 0.30; **38β**: 0.20; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **38α**: δ: 8.98 (s, br, 1H), 8.83 (s, 1H), 8.07 (s, 1H), 7.96 (d, J_1 = 8.0 Hz, 1H), 7.82 (d, J_1 = 8.2 Hz, 1H), 7.61 (td, J_1 = 6.8 Hz, J_2 = 1.1 Hz, 1H), 7.49 (td, J_1 = 6.9 Hz, J_2 = 1.1 Hz, 1H), 7.17 (t, J_1 = 7.8 Hz, 1H,

$H_{1'}$), 5.62 (ddd, J_1 = 8.6 Hz, J_2 = 3.6 Hz, J_3 = 2.3 Hz, 1H, $H_{3'}$), 4.55 (d, J_1 = 2.3 Hz, 1H, $H_{4'}$), 4.07 (dd, J_1 = 10.7 Hz, J_2 = 2.5 Hz, 1H, $H_{5'}$), 4.01 (dd, J_1 = 10.7 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 2.92-2.84 (m, 1H, $H_{2'}$), 2.65 (ddd, J_1 = 14.5 Hz, J_2 = 7.5 Hz, J_3 = 3.8 Hz, 1H, $H_{2'}$), 2.21 (s, 3H), 1.16-1.12 (m, 21H); **38β**: δ: 8.93 (s, br, 1H), 8.81 (s, 1H), 7.95 (d, J_1 = 8.2 Hz, 1H), 7.92 (s, 1H), 7.83 (d, J_1 = 8.2 Hz, 1H), 7.61 (td, J_1 = 6.9 Hz, J_2 = 1.2 Hz, 1H), 7.50 (td, J_1 = 6.9 Hz, J_2 = 1.1 Hz, 1H), 6.72 (t, J_1 = 7.4 Hz, 1H, $H_{1'}$), 5.57 (quint, J_1 = 3.9 Hz, 1H, $H_{3'}$), 4.20-4.16 (m, 1H, $H_{4'}$), 4.09-4.08 (m, 2H, $H_{5'}$ and $H_{5'}$), 3.32-3.24 (m, 1H, $H_{2'}$), 2.20 (ddd, J_1 = 14.2 Hz, J_2 = 7.4 Hz, J_3 = 3.7 Hz, 1H, $H_{2'}$), 2.12 (s, 3H), 1.15-1.05 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) **38α**: δ: 170.9, 161.8, 150.1, 136.6, 134.1, 130.9, 129.7, 129.5, 129.1, 127.6, 126.2, 116.6, 114.0, 86.6, 84.9, 76.2, 65.9, 34.9, 21.5, 18.2, 18.1, 12.1; **38β**: δ: 170.5, 161.9, 149.8, 136.7, 135.4, 131.0, 129.7, 129.4, 129.2, 127.9, 126.4, 116.3, 113.2, 85.2, 84.1, 74.1, 63.2, 34.5, 21.2, 18.1, 18.1, 12.1; **IR** (neat) **38α**: ν: 3191, 3060, 2941, 2865, 1739, 1709, 1684, 1632, 1602, 1474, 1230, 1122, 1013, 880, 743; **38β**: ν: 3192, 3055, 2942, 2864, 1737, 1691, 1632, 1602, 1476, 1384, 1234, 1065, 881, 731, 682 cm^{-1}; **HR-ESI MS** (m/z): $[M+Na]^+$ calcd for $C_{28}H_{38}N_2O_6SiNa$, 549.23968; **38α**: found 549.23897; **38β**: found 549.23871.

1'-(3'-*O*-Acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-α,β-D-ribofuranoside)-thieno[3,2-*d*]pyrimidine-2,4-(3H)-dione (39α) and **(39β)**



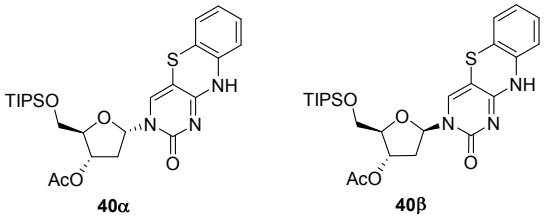
39

According to the general procedure A: The inseparable isomer mixture of nucleosides **39α** and **39β** (α/β = 1.0:2.0, 0.16 g, 73%) were obtained as off-white foams after column chromatography on silica gel (Hexane/EtOAc, 7:3 → 6:4) using thieno[3,2-*d*]pyrimidine-2,4-(3H)-dione **34** (0.092 g, 0.547 mmoles), BSA (0.28 mL, 1.14 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 0.20 g, 0.456 mmoles), NIS (0.113 g, 0.502 mmoles) and HOTf (16 μL, 0.182 mmoles).

Characterization: R_f (Hexane/EtOAc, 7:3): 0.16; **^{1}H NMR** (2D COSY, 500 MHz, $CDCl_3$) **39α**: δ: 9.45 (s, br, 1H), 7.73 (d, J_1 = 5.4 Hz, 1H), 7.39 (d, J_1 = 5.4 Hz, 1H), 6.87 (t, J_1 = 7.3 Hz, 1H, $H_{1'}$), 5.50-5.48 (m, 1H, $H_{3'}$), 4.37 (d, J_1 = 2.0 Hz, 1H, $H_{4'}$), 4.00 (dd, J_1 = 11.0 Hz, J_2 = 2.2 Hz, 1H, $H_{5'}$), 3.93 (dd, J_1 = 10.8 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 2.94-2.87 (m, 1H, $H_{2'}$), 2.34-2.28 (m, 1H, $H_{2'}$), 2.09 (s, 3H), 1.16-1.09 (m, 21H); **39β**: δ: 9.60 (s, br, 1H), 7.81 (d, J_1 = 5.4 Hz, 1H), 7.62 (d, J_1 = 5.4 Hz, 1H), 6.72 (dd, J_1 = 10.0 Hz, J_2 = 5.4

Hz, 1H, $H_{1'}$), 5.50-5.48 (m, 1H, $H_{3'}$), 4.12 (dd, J_1 = 11.3 Hz, J_2 = 2.3 Hz, 1H, $H_{5'}$), 4.06-3.98 (m, 2H, $H_{4'}$ and H_5), 2.56 (ddd, J_1 = 17.8 Hz, J_2 = 10.0 Hz, J_3 = 7.8 Hz, 1H, $H_{2'}$), 2.21 (dd, J_1 = 13.9 Hz, J_2 = 5.7 Hz, 1H, $H_{2'}$), 2.10 (s, 3H), 1.16-1.09 (m, 21H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 170.9, 170.7, 158.2, 158.0, 151.3, 151.0, 144.0, 134.7, 134.6, 120.2, 118.4, 115.6, 115.4, 86.6, 85.2, 84.8, 83.9, 75.6, 73.1, 65.6, 62.9, 36.7, 35.0, 21.2, 18.2, 18.1, 12.0; **IR** (neat) ν: 3187, 3057, 2942, 2865, 1740, 1698, 1486, 1236, 1220, 772 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{22}H_{34}N_2O_6SSiNa$, 505.18045; found 505.17942.

1'-(3'-*O*-Acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-α-D-ribofuranoside)-1,3-diaza-2-oxophenothiazine (40α) and **1'-(3'-*O*-acetyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-1,3-diaza-2-oxophenothiazine (40β)**

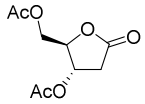


According to the general procedure A: The separable isomer mixture of nucleosides **40α** (0.040 g, 13%) and **40β** (0.100 g, 33%) were obtained as yellow solids after column chromatography on silica gel (Hexane/EtOAc, 4:6 → 0:1) using 1,3-diaza-2-oxophenothiazine **35** (0.148 g, 0.684 mmoles), BSA (0.35 mL, 1.425 mmoles), thioglycoside **14** (α/β = 1.0:1.8, 0.25 g, 0.570 mmoles), NIS (0.141 g, 0.627 mmoles) and HOTf (20 μL, 0.228 mmoles). Compound **40α** was precipitated in a mixture of Hexane/EtOAc.

Characterization: R_f (Hexane/EtOAc, 2:8): **40α**: 0.33; **40β**: 0.20; **^{1}H NMR** (2D COSY, 400 MHz, $CDCl_3$) **40α**: δ: 7.33 (s, 1H), 7.11-7.07 (m, 1H), 7.08-6.94 (m, 3H), 6.19 (d, J_1 = 6.4 Hz, 1H, $H_{1'}$), 5.32 (d, J_1 = 6.0 Hz, 1H, $H_{3'}$), 4.45 (t, J_1 = 2.8 Hz, 1H, $H_{4'}$), 3.92 (dd, J_1 = 11.0 Hz, J_2 = 2.6 Hz, 1H, $H_{5'}$), 3.84 (dd, J_1 = 11.0 Hz, J_2 = 3.4 Hz, 1H, $H_{5'}$), 2.88-2.81 (m, 1H, $H_{2'}$), 2.27 (d, J_1 = 15.3 Hz, 1H, $H_{2'}$), 2.02 (s, 3H), 1.09-1.05 (m, 21H); **40β**: δ: 7.57 (s, 1H), 7.24 (dd, J_1 = 8.0 Hz, J_2 = 1.0 Hz, 1H), 7.08-7.04 (m, 1H), 6.95-6.88 (m, 2H), 6.34 (dd, J_1 = 8.8 Hz, J_2 = 5.3 Hz, 1H, $H_{1'}$), 5.32 (d, J_1 = 6.2 Hz, 1H, $H_{3'}$), 4.11 (d, J_1 = 1.7 Hz, 1H, $H_{4'}$), 3.99 (d, J_1 = 2.0 Hz, 2H, $H_{5'}$ and $H_{5'}$), 2.59 (dd, J_1 = 13.9 Hz, J_2 = 5.0 Hz, 1H, $H_{2'}$), 2.10-2.03 (m, 1H, $H_{2'}$), 2.07 (s, 3H), 1.20-1.10 (m, 21H); **^{13}C NMR** (100 MHz, $CDCl_3$) **40α**: δ: 170.1, 158.2, 134.5, 134.2, 128.1, 126.3, 126.0, 118.6, 116.5, 96.0, 89.4, 88.9, 75.3, 64.3, 39.6, 21.2, 18.1, 12.0; **40β**: δ: 170.8, 160.8, 154.6, 135.9, 133.7, 127.7, 126.0, 124.6, 118.5, 116.7, 97.4, 86.5, 86.0, 75.5, 63.9, 39.1, 21.1, 18.2, 12.1; **IR**

(neat) **40α**: ν: 2942, 2865, 1741, 1655, 1470, 1444, 1415, 1254, 1129, 1094, 1013, 882, 751; **40β**: ν: 2941, 2864, 1741, 1654, 1623, 1593, 1435, 1419, 1232, 1192, 1120, 1006, 759, 682 cm^{-1}; **HR-ESI MS** (*m/z*): $[M+Na]^+$ calcd for $C_{26}H_{37}N_3O_5SSiNa$, 554.21209; **40α**: found 554.21139; **40β**: found 554.21159.

(4S,5R)-4-Acetoxy-5-acetoxymethyl-4,5-dihydrofuran-2-one (45)

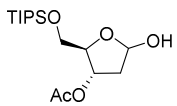


45

Procedure: To a solution of thioglycoside **11** (α/β = 1.0:1.8, 0.400 g, 1.233 mmoles) in EtOAc (32 mL), was added NBS (2.190 g, 12.330 mmoles) and H_2O (6.4 mL). The reaction mixture was stirred for 4h at room temperature and a sat. sol. $Na_2S_2O_3$ (25 mL) was added. The resulting solution was extracted with EtOAc (2 x 20 mL) and the combined organic layer was dried over $MgSO_4$, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane:EtOAc, 1:1) to give lactone **45** (0.162 g, 60%) as a colorless oil. The spectroscopic data are in good agreement with those previously reported.[310]

Characterization: $\mathbf{R_f}$ (Hexane/EtOAc, 4:6): 0.38; **^{1}H-NMR** (2D COSY, 500 MHz, $CDCl_3$) δ: 5.25 (dt, J_1 = 7.5 Hz, J_2 = 2.0 Hz, 1H, H_3), 4.65 (sext, J_1 = 1.8 Hz, 1H, H_4), 4.35 (dd, J_1 = 12.3 Hz, J_2 = 3.4 Hz, 1H, H_5), 4.25 (dd, J_1 = 12.3 Hz, J_2 = 3.6 Hz, 1H, H_5), 2.97 (dd, J_1 = 18.7 Hz, J_2 = 7.5 Hz, 1H, H_2), 2.58 (dd, J_1 = 18.7 Hz, J_2 = 2.1 Hz, 1H, H_2), 2.09 (s, 3H), 2.07 (s, 3H); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 173.9, 170.5, 170.2, 82.2, 71.2, 63.4, 34.9, 20.9, 20.8; **IR** (neat) ν: 2952, 2358, 1787, 1735, 1366, 1223, 1161, 1044, 939 cm^{-1}; **HR-ESI MS** (*m/z*): $[M+Na]^+$ calcd for $C_9H_{12}O_6Na$, 239.05316; found 239.05264.

3-*O*-acetyl-5-*O*-triisopropylsilyl-2-deoxy-α,β-D-ribofuranoside (46)



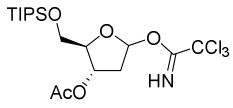
46

Procedure: To a solution of thioglycoside **14** (α/β = 1.0:1.8, 0.320 g, 0.729 mmoles) in a mixture of CH_3CN/H_2O (10:1, 4.07 mL) at room temperature, was added *N*-Iodosaccharin (NISac, 0.083 g, 0.37 mmoles). The reaction mixture was stirred for 10 min and quenched with aq. sat. sol. $NaHCO_3$. It was then extracted with CH_2Cl_2 and the combined organics layers were dried over $MgSO_4$, filtered and

evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (Hexane/EtOAc, 2:8) to afford lactol **46** (a/b = 1.2:1.0, 0.175 g, 72%) as a colorless oil.

Characterization: R_f (Hexane/EtOAc, 7:3): 0.26; **^{1}H-NMR** (2D COSY, 400 MHz, $CDCl_3$) **46a**: δ: 5.50 (q, J_1 = 2.4 Hz, 1H, H_1), 5.34 (ddd, J_1 = 7.0 Hz, J_2 = 4.4 Hz, J_3 = 2.2 Hz, 1H, H_3), 4.10 (q, J_1 = 2.7 Hz, 1H, H_4), 3.94 (dd, J_1 = 10.6 Hz, J_2 = 2.7 Hz, 1H, H_5), 3.79 (dd, J_1 = 10.6 Hz, J_2 = 3.2 Hz, 1H, H_5), 2.24 (m, 2H, H_2), 2.01 (s, 3H), 1.09 (s, 3H), 1.03 (m, 18H); **46b**: δ: 5.54 (d, J_1 = 4.8 Hz, 1H, H_1), 5.23 (dq, J_1 = 7.1 Hz, J_2 = 1.3 Hz, 1H, H_3), 4.24 (q, J_1 = 3.4 Hz, 1H, H_4), 3.84 (dd, J_1 = 10.8 Hz, J_2 = 3.2 Hz, 1H, H_5), 3.71 (dd, J_1 = 11.0 Hz, J_2 = 3.6 Hz, 1H, H_5), 2.36 - 2.19 (m, 2H, H_2), 2.05 (s, 3H), 1.09 (s, 3H), 1.03 (m, 18H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 171.02, 170.86, 99.26, 99.21, 85.39, 85.09, 75.63, 75.58, 64.80, 64.11, 41.91, 40.10, 21.27, 21.09, 18.03, 18.00, 17.99, 17.98, 12.00, 11.96; **IR** (neat) ν: 3422, 2942, 2866, 2359, 2341, 1738, 1240, 1062, 881, 680, 669, 419, 408 cm^{-1}.

1-*O*-trichloroacetimidate-3-*O*-acetyl-5-*O*-triisopropylsilyl-2-deoxy-α,β-D-ribofuranoside (47)



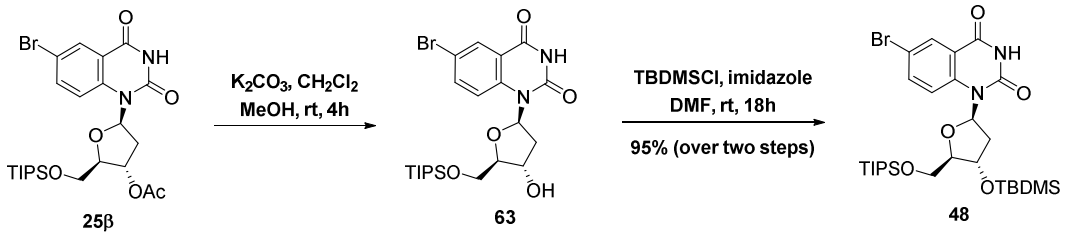
47

Procedure: To a solution of lactol **46** (0.009 g, 0.027 mmoles) in CH_2Cl_2 (0.5 mL), was added CCl_3CN (0.019 mL, 0.27 mmoles) and NaH (60% dispersion in mineral oil, 0.0022 g, 0.068 mmoles). The reaction mixture was stirred at room temperature for 2h. The solids were then filtered off using a fritted glass and the filtrate was evaporated to give crude trichloroacetimidate **47** as a colorless oil which was used directly in the next step without further purification due to its instability on silica gel. The spectroscopic data are reported for the major diastereoisomer.

Characterization: **^{1}H-NMR** (2D COSY, 400 MHz, $CDCl_3$) δ: 8.41 (s, br, 1H), 6.48 (d, J_1 = 4.7 Hz, 1H, H_1), 5.33 (d, J_1 = 7.2 Hz, 1H, H_3), 4.40 (d, J_1 = 1.2 Hz, 1H, H_4), 3.97 (dd, J_1 = 10.9 Hz, J_2 = 2.6 Hz, 1H, H_5), 3.86 (dd, J_1 = 10.9 Hz, J_2 = 3.3 Hz, 1H, H_5), 2.48 (m, 1H, H_2), 2.34 (m, 1H, H_2), 2.05 (s, 3H), 1.08-1.05 (m, 21H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 171.21, 161.60, 103.48, 91.81, 87.69, 74.41, 63.68, 38.97, 21.26, 18.15, 18.14, 18.12, 18.11, 12.10; **IR** (neat) ν: 2942, 2925, 2867, 2360, 2341, 1738, 1238, 1106, 1066, 882, 798, 682 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M+Na]^+$ calcd for $C_{18}H_{32}Cl_3NO_5SiNa$, 498.10131; found

498.10061. The presence of three chlorines was observed with the isotopic pattern of the molecular ions, found 500.09789, 502.09523 and 504.09368.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-bromo-quinazoline-2,4-(3*H*)-dione (48)

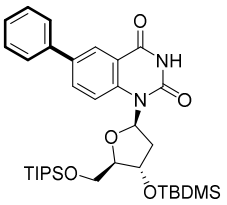


Procedure: To a stirred solution of **25β** (6.015 g, 10.827 mmol) in a mixture of MeOH/CH_2Cl_2 (75 mL, 4:1), was added K_2CO_3 (1.795 g, 12.993 mmoles). The reaction mixture was stirred at room temperature for 18h and the solvents were evaporated. The crude material was taken up in EtOAc (200 mL) and washed with sat. sol. NaCl (200 mL). The aqueous phase was back extracted with EtOAc (2 x 50 mL) and the combined organic layers was dried, filtered and evaporated *in vacuo* to give the deacetylated nucleoside **63** as an off-white foam. The product was used in the next step without further purification. $\mathbf{R}_f$ (CH_2Cl_2/MeOH, 96:4): 0.21. To a stirred solution of **63** (5.559 g, 10.827 mmoles) in DMF (85 mL), was added TBDMSCl (1.958 g, 12.993 mmoles) and imidazole (1.105 g, 16.241 mmoles). The reaction mixture was stirred at room temperature for 18h and then quenched with sat. sol. NaCl (40 mL). The resulting solution was extracted with EtOAc (3 x 40 mL) and the combined organic layers were dried, filtered and evaporated *in vacuo*. The crude material was subjected to column chromatography on silica gel (hexane/EtOAc, 93:7 → 90:10) to give nucleoside **48** (6.443 g, 95% over two steps from **25β**) as a white foam.

Characterization: $\mathbf{R}_f$ (hexane/EtOAc, 9:1): 0.16; $[\alpha]_D^{25}$ +35.6 (*c* 1.0, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.61 (s, 1H), 8.31 (d, J_1 = 2.4 Hz, 1H), 7.68 (d, J_1 = 9.1 Hz, 1H), 7.62 (dd, J_1 = 9.1 Hz, J_2 = 2.4 Hz, 1H), 6.74 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 4.72 (quint, J_1 = 3.9 Hz, 1H, $H_{3'}$), 4.03 (dd, J_1 = 11.4 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.91 (dd, J_1 = 11.4 Hz, J_2 = 3.2 Hz, 1H, $H_{5'}$), 3.85-3.82 (m, 1H, $H_{4'}$), 2.75-2.68 (m, 1H, $H_{2'}$), 2.05 (ddd, J_1 = 13.4 Hz, J_2 = 7.4 Hz, J_3 = 3.7 Hz, 1H, $H_{2'}$), 1.16-1.07 (m, 21H), 0.90 (s, 9H), 0.09 (s, 3H), 0.08 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 161.0, 150.0, 138.9, 137.7, 131.2, 119.1, 118.6, 117.0, 86.6, 84.2, 70.1, 61.9, 37.1, 25.9, 18.2, 12.1, -4.4, -4.7; **IR** (neat) ν: 3197, 3073, 2944, 2896, 2865, 1714, 1603, 1484,

1465, 1360, 1316, 1252, 1110, 1061, 882, 836, 778, 681 cm^{-1}; **HR-ESI MS** (*m/z*): $[M + Na]^+$ calcd for $C_{28}H_{47}BrN_2O_5Si_2Na$ 651.20841, found 651.20776.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-phenyl-quinazoline-2,4-(3*H*)-dione (49a)

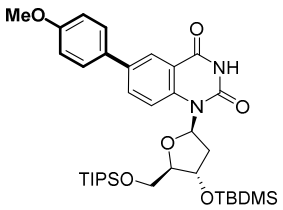


49a

According to the general procedure B: Biaryl nucleoside **49a** (0.361 g, 91%) was obtained as a white foam after column chromatography on silica gel (hexane/EtOAc, 90:10 → 85:15) using nucleoside **48** (0.400 g, 0.637 mmol), phenylboronic acid (0.117 g, 0.956 mmol), AcOK (0.094 g, 0.956 mmol) and Pd(dppf)Cl_2 (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.31; $[\alpha]_D^{26}$ +30.3 (*c* 0.5, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.56 (s, 1H), 8.46 (s, 1H), 7.85 (s, 2H), 7.62 (d, J_1 = 7.6 Hz, 2H), 7.47 (t, J_1 = 7.6 Hz, 2H), 7.40-7.36 (m, 1H), 6.82 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 4.78 (quint, J_1 = 4.0 Hz, 1H, $H_{3'}$), 4.08 (dd, J_1 = 11.2 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.96 (dd, J_1 = 11.2 Hz, J_2 = 3.2 Hz, 1H, $H_{5'}$), 3.90-3.87 (m, 1H, $H_{4'}$), 2.88-2.81 (m, 1H, $H_{2'}$), 2.10 (ddd, J_1 = 13.2 Hz, J_2 = 7.6 Hz, J_3 = 3.6 Hz, 1H, $H_{2'}$), 1.19-1.09 (m, 21H), 0.92 (s, 9H), 0.12 (s, 3H), 0.11 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.2, 150.2, 139.1, 139.0, 136.7, 133.4, 129.2, 128.0, 126.9, 126.7, 117.7, 117.3, 86.6, 84.2, 70.4, 62.1, 37.2, 26.0, 18.2, 12.2, -4.4, -4.6; **IR** (neat) ν: 3192, 3067, 2944, 2864, 1688, 1620, 1471, 1367, 1320, 1251, 1109, 1065, 1050, 881, 834, 777, 759, 731, 676 cm^{-1}; **HR-ESI MS** (*m/z*): $[M + Na]^+$ calcd for $C_{34}H_{52}N_2O_5Si_2Na$ 647.33125, found 647.33145.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(4-methoxy)-phenyl-quinazoline-2,4-(3*H*)-dione (49b)

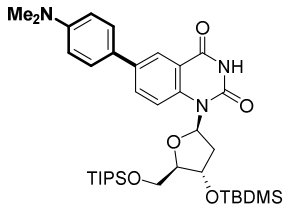


49b

According to the general procedure B: Biaryl nucleoside **49b** (0.377 g, 90%) was obtained as a white foam after column chromatography on silica gel (hexane/EtOAc, 85:15 → 80:20) using nucleoside **48** (0.400 g, 0.637 mmol), 4-methoxyphenylboronic acid (0.145 g, 0.956 mmol), AcOK (0.094 g, 0.956 mmol) and Pd(dppf)Cl_2 (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.33; $[\alpha]_D^{27}$ +31.4 (*c* 1.0, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ: 9.68 (s, 1H), 8.41 (d, J_1 = 2.3 Hz, 1H), 7.84 (d, J_1 = 8.9 Hz, 1H), 7.78 (dd, J_1 = 8.9 Hz, J_2 = 2.3 Hz, 1H), 7.56 (d, J_1 = 6.7 Hz, 2H), 7.02 (d, J_1 = 9.8 Hz, 2H), 6.86 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 4.80-4.77 (m, 1H, $H_{3'}$), 4.07 (dd, J_1 = 11.4 Hz, J_2 = 2.6 Hz, 1H, $H_{5'}$), 3.95 (dd, J_1 = 11.4 Hz, J_2 = 3.1 Hz, 1H, $H_{5'}$), 3.89-3.86 (m, 4H, $H_{4'}$), 2.82-2.76 (m, 1H, $H_{2'}$), 2.08 (ddd, J_1 = 13.4 Hz, J_2 = 7.6 Hz, J_3 = 3.8 Hz, 1H, $H_{2'}$), 1.19-1.08 (m, 21H), 0.90 (s, 9H), 0.10 (s, 3H), 0.09 (s, 3H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ: 162.3, 159.6, 150.3, 138.3, 136.3, 133.0, 131.3, 128.0, 126.0, 117.7, 117.2, 114.6, 86.3, 83.8, 69.9, 61.7, 55.5, 37.0, 25.9, 18.2, 12.0, -4.4, -4.7; **IR** (neat) ν: 3188, 3067, 2942, 2864, 1687, 1619, 1495, 1468, 1368, 1321, 1250, 1110, 1049, 835, 778, 732, 678 cm^{-1}; **HR-ESI MS** (*m/z*): $[M + Na]^+$ calcd for $C_{35}H_{54}N_2O_6Si_2Na$ 677.34126, found 677.34043.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(4-*N*,*N*-dimethylamino)-phenyl-quinazoline-2,4-(3*H*)-dione (49c)



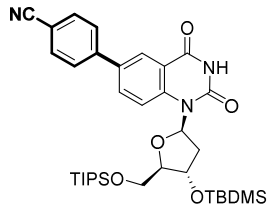
49c

According to the general procedure B: Biaryl nucleoside **49c** (2.852 g, 67%) was obtained as a yellowish foam after column chromatography on silica gel (Hexane/EtOAc, 80:20 → 70:30) using nucleoside **48** (4.00 g, 6.372 mmol), 4-(*N*,*N*-dimethylamino)phenylboronic acid (2.313 g, 14.018 mmol), AcOK (1.250 g, 12.744 mmol) and Pd(dppf)Cl_2 (0.260 g, 0.319 mmol).

Characterization: R_f (hexane/EtOAc, 7:3): 0.23; $[\alpha]_D^{25}$ +32.8 (*c* 0.5, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.57 (s, 1H), 8.41 (s, 1H), 7.81 (s, 2H), 7.54 (d, J_1 = 8.8 Hz, 2H), 6.84-6.81 (m, 3H, $H_{1'}$), 4.79 (quint, J_1 = 4.4 Hz, 1H, $H_{3'}$), 4.08 (dd, J_1 = 11.6 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.97 (dd, J_1 = 11.6 Hz, J_2 = 3.6 Hz, 1H, $H_{5'}$), 3.89-3.87 (m, 1H, $H_{4'}$), 3.02 (s, 6H), 2.89-2.81 (m, 1H, $H_{2'}$), 2.09 (ddd, J_1 = 13.2 Hz, J_2 = 7.6 Hz,

J_3 = 3.6 Hz, 1H, $H_{2'}$), 1.18-1.10 (m, 21H), 0.93 (s, 9H), 0.12 (s, 3H), 0.11 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.4, 150.36, 150.28, 137.9, 136.7, 132.5, 127.5, 126.8, 125.2, 117.5, 117.2, 113.0, 86.5, 84.1, 70.3, 62.1, 40.7, 37.1, 25.9, 18.2, 12.1, -4.4, -4.7; **IR** (neat) ν: 3193, 3062, 2944, 2864, 1686, 1610, 1497, 1467, 1362, 1321, 1109, 1064, 835, 777, 729 cm^{-1}; **HR-ESI MS** (*m*/*z*): [M + Na]$^+$ calcd for $C_{36}H_{57}N_3O_5Si_2Na$ 690.37345, found 690.37390.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(4-cyano)-phenyl-quinazoline-2,4-(3*H*)-dione (49d)

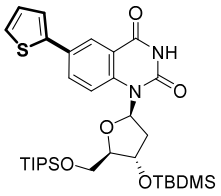


49d

According to the general procedure A: Biaryl nucleoside **49d** (0.378 g, 91%) was obtained as a white foam after column chromatography on silica gel (hexane/EtOAc, 85:15 → 80:20) using nucleoside **48** (0.400 g, 0.637 mmol), 4-cyanophenylboronic acid (0.126 g, 0.956 mmol), AcOK (0.094 g, 0.956 mmol) and Pd(dppf)Cl_2 (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.19; $[\alpha]_D^{27}$ +29.7 (*c* 1.0, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.82 (s, 1H), 8.46 (d, J_1 = 2.4 Hz, 1H), 7.90 (d, J_1 = 8.8 Hz, 1H), 7.81 (dd, J_1 = 9.2 Hz, J_2 = 2.4 Hz, 1H), 7.74 (dd, J_1 = 18.0 Hz, J_2 = 8.4 Hz, 4H), 6.81 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 4.77 (quint, J_1 = 4.4 Hz, 1H, $H_{3'}$), 4.06 (dd, J_1 = 11.6 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.95 (dd, J_1 = 11.2 Hz, J_2 = 3.2 Hz, 1H, $H_{5'}$), 3.89-3.87 (m, 1H, $H_{4'}$), 2.85-2.78 (m, 1H, $H_{2'}$), 2.10 (ddd, J_1 = 13.2 Hz, J_2 = 7.6 Hz, J_3 = 4.0 Hz, 1H, $H_{2'}$), 1.18-1.10 (m, 21H), 0.93 (s, 9H), 0.12 (s, 3H), 0.11 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.0, 150.1, 143.3, 140.2, 134.4, 133.2, 133.0, 127.5, 127.1, 118.8, 118.1, 117.5, 111.7, 86.7, 84.2, 70.3, 62.1, 37.2, 25.9, 18.2, 12.1, -4.4, -4.7; **IR** (neat) ν: 3192, 3067, 2944, 2864, 2228, 1690, 1619, 1490, 1468, 1109, 1063, 909, 834, 778, 731 cm^{-1}; **HR-ESI MS** (*m*/*z*): [M + Na]$^+$ calcd for $C_{35}H_{51}N_3O_5Si_2Na$ 672.32650, found 672.32632.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(2-thiophene)-quinazoline-2,4-(3*H*)-dione (49e)

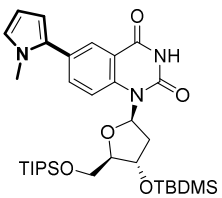


49e

According to the general procedure A: Biaryl nucleoside **49e** (0.249 g, 62%) was obtained as a white foam after column chromatography on silica gel (hexane/EtOAc, 85:15 → 80:20) using nucleoside **48** (0.400 g, 0.637 mmol), 2-thienylboronic acid (0.122 g, 0.956 mmol), AcOK (0.094 g, 0.956 mmol) and Pd(dppf)Cl_2 (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.33; $[\alpha]_D^{27}$ +35.1 (*c* 1.0, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.40 (s, 1H), 8.42 (t, J_1 = 1.3 Hz, 1H), 7.81 (d, J_1 = 1.4 Hz, 2H), 7.37 (dd, J_1 = 3.6 Hz, J_2 = 1.1 Hz, 1H), 7.31 (dd, J_1 = 5.1 Hz, J_2 = 1.1 Hz, 1H), 7.11 (dd, J_1 = 5.1 Hz, J_2 = 3.6 Hz, 1H), 6.81 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 4.80-4.76 (m, 1H, $H_{3'}$), 4.07 (dd, J_1 = 11.3 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.95 (dd, J_1 = 11.4 Hz, J_2 = 3.3 Hz, 1H, $H_{5'}$), 3.88-3.85 (m, 1H, $H_{4'}$), 2.83-2.76 (m, 1H, $H_{2'}$), 2.09 (ddd, J_1 = 13.4 Hz, J_2 = 7.5 Hz, J_3 = 3.8 Hz, 1H, $H_{2'}$), 1.17-1.09 (m, 21H), 0.92 (s, 9H), 0.11 (s, 3H), 0.10 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 161.9, 150.1, 142.3, 138.8, 132.3, 130.4, 128.5, 125.6, 125.1, 123.8, 117.8, 117.3, 86.6, 84.1, 70.2, 62.0, 37.2, 25.9, 18.2, 12.2, -4.4, -4.7; **IR** (neat) ν: 3188, 3067, 2940, 2864, 1690, 1495, 1466, 1372, 1312, 1251, 1109, 1062, 882, 835, 777, 731, 689 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{32}H_{50}N_2O_5SSi_2Na$ 653.28712, found 653.28670.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(*N*-methyl-2-pyrrole)-quinazoline-2,4-(3*H*)-dione (49f)



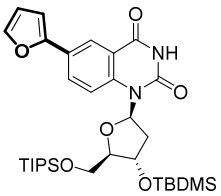
49f

According to the general procedure A: Biaryl nucleoside **49f** (0.327 g, 82%) was obtained as a white foam after column chromatography on silica gel (hexane/EtOAc, 80:20 → 70:30) using nucleoside

48 (0.400 g, 0.637 mmol), *N*-methyl-2-pyrroleboronic acid pinacol ester (0.330 g, 1.593 mmol), AcOK (0.156 g, 1.593 mmol) and Pd(dppf)Cl_2 (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.24; $[\alpha]_D^{26}$ +38.9 (*c* 1.0, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.54 (s, 1H), 8.24 (d, J_1 = 2.2 Hz, 1H), 7.84 (d, J_1 = 8.8 Hz, 1H), 7.65 (dd, J_1 = 8.8 Hz, J_2 = 2.3 Hz, 1H), 6.83 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 6.75 (dd, J_1 = 2.4 Hz, J_2 = 2.0 Hz, 1H), 6.29 (dd, J_1 = 3.6 Hz, J_2 = 1.8 Hz, 1H), 6.22 (dd, J_1 = 3.6 Hz, J_2 = 2.7 Hz, 1H), 4.80-4.76 (m, 1H, $H_{3'}$), 4.08 (dd, J_1 = 11.4 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.97 (dd, J_1 = 11.4 Hz, J_2 = 3.2 Hz, 1H, $H_{5'}$), 3.88-3.87 (m, 1H, $H_{4'}$), 3.71 (s, 3H), 2.86-2.78 (m, 1H, $H_{2'}$), 2.09 (ddd, J_1 = 13.4 Hz, J_2 = 7.4 Hz, J_3 = 3.7 Hz, 1H, $H_{2'}$), 1.23-1.11 (m, 21H), 0.92 (s, 9H), 0.11 (s, 3H), 0.10 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.0, 150.2, 138.3, 134.7, 132.5, 129.1, 127.5, 124.7, 117.4, 117.0, 109.5, 108.3, 86.6, 84.1, 70.2, 62.0, 37.1, 35.4, 25.9, 18.2, 12.1, -4.4, -4.7; **IR** (neat) ν: 3188, 3067, 2944, 2864, 1687, 1500, 1459, 1318, 1109, 1057, 834, 777, 729, 509 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{33}H_{53}N_3O_5Si_2Na$ 650.34160, found 650.34150.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(2-furan)-quinazoline-2,4-(3*H*)-dione (49g)



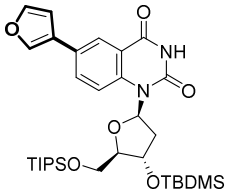
49g

According to the general procedure A: Biaryl nucleoside **49g** (0.252 g, 64%) was obtained as a yellowish foam after column chromatography on silica gel (hexane/EtOAc, 90:10 → 80:20) using nucleoside **48** (0.400 g, 0.637 mmol), 2-furanboronic acid (0.107 g, 0.956 mmol), AcOK (0.094 g, 0.956 mmol) and Pd(dppf)Cl_2 (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.24; $[\alpha]_D^{26}$ +34.0 (*c* 1.0, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.55 (s, 1H), 8.46 (d, J_1 = 2.2 Hz, 1H), 7.88 (dd, J_1 = 8.9 Hz, J_2 = 2.2 Hz, 1H), 7.78 (d, J_1 = 9.0 Hz, 1H), 7.49 (d, J_1 = 1.2 Hz, 1H), 6.78 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 6.71 (d, J_1 = 3.2 Hz, 1H), 6.49 (dd, J_1 = 3.4 Hz, J_2 = 1.8 Hz, 1H), 4.77 (quint, J_1 = 4.7 Hz, 1H, $H_{3'}$), 4.06 (dd, J_1 = 11.3 Hz, J_2 = 3.0 Hz, 1H, $H_{5'}$), 3.94 (dd, J_1 = 11.4 Hz, J_2 = 3.5 Hz, 1H, $H_{5'}$), 3.88-3.86 (m, 1H, $H_{4'}$), 2.85-2.78 (m, 1H, $H_{2'}$), 2.08 (ddd, J_1 = 13.4 Hz, J_2 = 7.6 Hz, J_3 = 3.9 Hz, 1H, $H_{2'}$), 1.16-1.08 (m, 21H), 0.92 (s, 9H), 0.11 (s, 3H), 0.10 (s, 3H); **^{13}C**

NMR (100 MHz, $CDCl_3$) δ: 162.0, 152.3, 150.1, 142.8, 138.8, 130.1, 126.9, 123.5, 117.6, 117.3, 112.1, 106.0, 86.7, 84.2, 70.4, 62.1, 37.2, 25.9, 18.2, 12.2, -4.4, -4.7; **IR** (neat) ν: 3188, 3067, 2944, 2864, 1690, 1461, 1315, 1109, 1065, 1013, 883, 834, 777, 731, 512 cm^{-1}; **HR-ESI MS** (m/z): $[M + Na]^+$ calcd for $C_{32}H_{50}N_2O_6Si_2Na$ 637.30996, found 637.30968.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(3-furan)-quinazoline-2,4-(3*H*)-dione (49h)

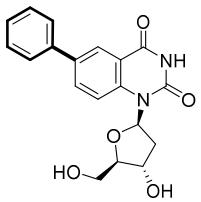


49h

According to the general procedure A: Biaryl nucleoside **49h** (0.273 g, 70%) was obtained as a white foam after column chromatography on silica gel (hexane/EtOAc, 90:10 → 80:20) using nucleoside **48** (0.400 g, 0.637 mmol), 3-furanboronic acid pinacol ester (0.186 g, 0.956 mmol), AcOK (0.094 g, 0.956 mmol) and $Pd(dppf)Cl_2$ (0.026 g, 0.032 mmol).

Characterization: R_f (hexane/EtOAc, 8:2): 0.26; $[\alpha]_D^{26}$ +35.5 (*c* 0.5, $CHCl_3$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.66 (s, 1H), 8.30 (d, J_1 = 2.0 Hz, 1H), 7.80-7.77 (m, 2H), 7.68 (dd, J_1 = 8.8 Hz, J_2 = 2.0 Hz, 1H), 7.50 (t, J_1 = 1.6 Hz, 1H), 6.79 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 6.73 (dd, J_1 = 1.6 Hz, J_1 = 0.8 Hz, 1H), 4.78-4.74 (m, 1H, $H_{3'}$), 4.06 (dd, J_1 = 11.6 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.94 (dd, J_1 = 11.2 Hz, J_2 = 3.2 Hz, 1H, $H_{5'}$), 3.88-3.85 (m, 1H, $H_{4'}$), 2.84-2.77 (m, 1H, $H_{2'}$), 2.08 (ddd, J_1 = 13.2 Hz, J_2 = 7.6 Hz, J_3 = 4.0 Hz, 1H, $H_{2'}$), 1.17-1.10 (m, 21H), 0.91 (s, 9H), 0.11 (s, 3H), 0.09 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.2, 150.2, 144.3, 139.0, 138.7, 132.2, 128.4, 125.3, 124.9, 117.7, 117.3, 108.7, 86.6, 84.1, 70.3, 62.0, 37.2, 25.9, 18.2, 12.2, -4.4, -4.7; **IR** (neat) ν: 3193, 3067, 2942, 2864, 1689, 1469, 1320, 1108, 1058, 874, 834, 778, 730, 512 cm^{-1}; **HR-ESI MS** (m/z): $[M + Na]^+$ calcd for $C_{32}H_{50}N_2O_6Si_2Na$ 637.30996, found 637.31006.

1'-(2'-deoxy-β-D-ribofuranoside)-6-phenyl-quinazoline-2,4-(3*H*)-dione (50a)

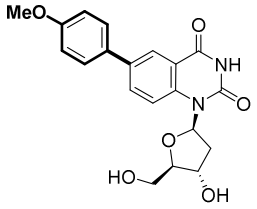


50a

According to the general procedure C: Emissive nucleoside **50a** (0.040 g, 70%) was obtained as a white solid after column chromatography on silica gel (MeCN/MeOH, 99:1) using biaryl nucleoside **49a** (0.100 g, 0.160 mmol) and TBAF (0.167 g, 0.640 mmol)

Characterization: R_f (CH_2Cl_2/MeOH, 95:5): 0.16; $[\alpha]_D^{27}$ +28.7 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.69 (s, br, 1H), 8.23 (s, 1H), 7.96 (d, J_1 = 1.2 Hz, 2H), 7.72 (d, J_1 = 7.2 Hz, 2H), 7.49 (t, J_1 = 7.2 Hz, 2H), 7.39 (t, J_1 = 7.2 Hz, 1H), 6.71 (t, J_1 = 8.0 Hz, 1H, $H_{1'}$), 5.28 (d, J_1 = 4.4 Hz, 1H), 4.99 (s, 1H), 4.43-4.41 (m, 1H, $H_{3'}$), 3.75-3.64 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.72-2.64 (m, 1H, $H_{2'}$), 1.96 (ddd, J_1 = 13.6 Hz, J_2 = 7.2 Hz, J_3 = 3.2 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.5, 149.9, 138.7, 138.2, 134.6, 132.5, 129.1, 127.7, 126.3, 124.7, 117.9, 117.0, 86.5, 83.4, 69.6, 60.9, 36.2; **IR** (neat) ν: 3468, 3312, 2974, 2921, 2817, 1685, 1618, 1474, 1327, 1112, 1078, 997, 943, 762, 513 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 269 (2.5 x 10^4); **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{19}H_{18}N_2O_5Na$ 377.11079, found 377.11054.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(4-methoxy)-phenyl-quinazoline-2,4-(3*H*)-dione (50b)

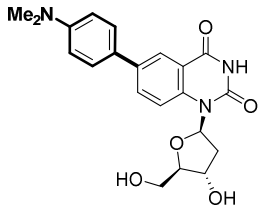


50b

According to the general procedure C: Emissive nucleoside **50b** (0.053 g, 90%) was obtained as a white solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5 → 90:10) using biaryl nucleoside **49b** (0.100 g, 0.153 mmol) and TBAF (0.160 g, 0.611 mmol).

Characterization: R_f (CH_2Cl_2/MeOH, 95:5): 0.21; $[\alpha]_D^{26}$ +30.9 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.66 (s, br, 1H), 8.17 (s, 1H), 7.93-7.91 (m, 2H), 7.65 (d, J_1 = 8.5 Hz, 2H), 7.03 (d, J_1 = 8.4 Hz, 2H), 6.70 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 5.27 (s, 1H), 4.98 (s, 1H), 4.42 (s, 1H, $H_{3'}$), 3.80 (s, 3H), 3.74-3.64 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.71-2.63 (m, 1H, $H_{2'}$), 1.98-1.93 (m, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.6, 159.1, 149.9, 138.1, 134.5, 132.1, 130.5, 127.5, 124.0, 117.8, 116.9, 114.6, 86.5, 83.4, 69.6, 60.9, 55.2, 36.2; **IR** (neat) ν: 3252, 2835, 1697, 1676, 1495, 1469, 1320, 1247, 1175, 1097, 1058, 1020, 816, 806, 508 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 276 (2.7 x 10^4); **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{20}H_{20}N_2O_6Na$ 407.12136, found 407.12099.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(4-*N*,*N*-dimethylamino)-phenyl-quinazoline-2,4-(3*H*)-dione (50c)

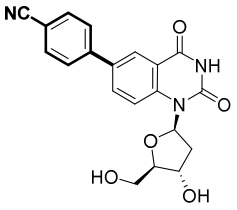


50c

According to the general procedure C: Emissive nucleoside **50c** (0.0474 g, 80%) was obtained as a yellow solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5 → 85:15) using biaryl nucleoside **49c** (0.100 g, 0.149 mmol) and TBAF (0.157 g, 0.599 mmol).

Characterization: R_f (CH_2Cl_2/MeOH, 92:8): 0.25; $[\alpha]_D^{26}$ + 32.2 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.66 (s, br, 1H), 8.14 (s, 1H), 7.89 (s, 2H), 7.55 (d, J_1 = 8.7 Hz, 2H), 6.81 (d, J_1 = 8.8 Hz, 2H), 6.69 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.26 (d, J_1 = 4.9 Hz, 1H), 4.97 (t, J_1 = 5.0 Hz, 1H), 4.41-4.39 (m, 1H, $H_{3'}$), 3.73-3.64 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.95 (s, 6H), 2.70-2.63 (m, 1H, $H_{2'}$), 1.94 (ddd, J_1 = 13.4 Hz, J_2 = 7.0 Hz, J_3 = 3.4 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.7, 150.0, 149.9, 137.4, 135.0, 131.4, 126.8, 125.5, 123.0, 117.7, 116.9, 112.7, 86.5, 83.3, 69.6, 60.9, 39.9, 36.2; **IR** (neat) ν: 3468, 3312, 2974, 2818, 1683, 1607, 1496, 1472, 1325, 1313, 1207, 1110, 997, 941, 808, 509 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 297 (2.0 x 10^4); **HR-ESI MS** (*m*/*z*): $[M + H]^+$ calcd for $C_{21}H_{24}N_3O_5$ 398.17105, found 398.17144.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(4-cyano)-phenyl-quinazoline-2,4-(3*H*)-dione (50d)

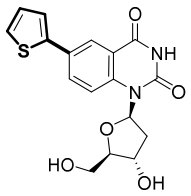


50d

According to the general procedure C: Emissive nucleoside **50d** (0.0537 g, 92%) was obtained as a white solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5) using biaryl nucleoside **49d** (0.100 g, 0.154 mmol) and TBAF (0.161 g, 0.615 mmol).

Characterization: R_f (CH_2Cl_2/MeOH, 95:5): 0.17; $[\alpha]_D^{27}$ +40.5 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.74 (s, br, 1H), 8.31 (d, J_1 = 2.1 Hz, 1H), 8.05 (dd, J_1 = 9.0 Hz, J_2 = 2.2 Hz, 1H), 8.00 (d, J_1 = 8.9 Hz, 1H), 7.96-7.91 (m, 4H), 6.71 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 5.28 (d, J_1 = 4.8 Hz, 1H), 4.98 (s, 1H), 4.43-4.41 (m, 1H, $H_{3'}$), 3.74-3.64 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.70-2.63 (m, 1H, $H_{2'}$), 1.96 (ddd, J_1 = 13.3 Hz, J_2 = 7.4 Hz, J_3 = 3.5 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.4, 149.9, 142.6, 139.6, 133.0, 132.7, 132.6, 127.2, 125.4, 118.7, 118.1, 117.1, 110.2, 86.5, 83.5, 69.6, 60.9, 36.3; **IR** (neat) ν: 3479, 3405, 3047, 3854, 2223, 1680, 1618, 1603, 1471, 1327, 1082, 820, 544 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 286 (2.1 x 10^4); **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{20}H_{17}N_3O_5Na$ 402.10659, found 402.10640.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(2-thiophene)-quinazoline-2,4-(3*H*)-dione (50e)

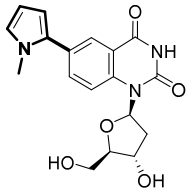


50e

According to the general procedure C: Emissive nucleoside **50e** (0.040 g, 70%) was obtained as a white solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5 → 85:15) using biaryl nucleoside **49e** (0.100 g, 0.158 mmol) and TBAF (0.166 g, 0.634 mmol). A precipitate was formed during the reaction.

Characterization: R_f (CH_2Cl_2/MeOH, 92:8): 0.26; $[\alpha]_D^{26}$ +36.3 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.71 (s, br, 1H), 8.16 (t, J_1 = 1.5 Hz, 1H), 7.94-7.93 (m, 2H), 7.60-7.57 (m, 2H), 7.15 (dd, J_1 = 5.0 Hz, J_2 = 3.7 Hz, 1H), 6.69 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.26 (d, J_1 = 4.7 Hz, 1H), 4.99 (s, 1H), 4.42-4.40 (m, 1H, $H_{3'}$), 3.74-3.67 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.68-2.60 (m, 1H, $H_{2'}$), 1.95 (ddd, J_1 = 13.4 Hz, J_2 = 7.3 Hz, J_3 = 3.5 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.4, 149.9, 141.5, 138.5, 131.1, 128.8, 128.7, 126.0, 124.1, 123.2, 118.1, 117.0, 86.5, 83.4, 69.5, 60.8, 36.2; **IR** (neat) ν: 3400, 3333, 3166, 3051, 2896, 1714, 1697, 1655, 1496, 1470, 1084, 1035, 1005, 809, 719, 710, 512, 503 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 295 (2.0 x 10^4); **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{17}H_{16}N_2O_5SNa$ 383.06721, found 383.06672.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(*N*-methyl-2-pyrrole)-quinazoline-2,4-(3*H*)-dione (50f)

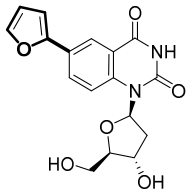


50f

According to the general procedure C: Emissive nucleoside **50f** (0.046 g, 81%) was obtained as a white solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5 → 90:10) using biaryl nucleoside **49f** (0.100 g, 0.159 mmol) and TBAF (0.166 g, 0.637 mmol).

Characterization: R_f (CH_2Cl_2/MeOH, 95:5): 0.16; $[\alpha]_D^{26}$ +35.2 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.66 (s, br, 1H), 7.98 (d, J_1 = 2.2 Hz, 1H), 7.89 (d, J_1 = 8.9 Hz, 1H), 7.71 (dd, J_1 = 8.8 Hz, J_2 = 2.2 Hz, 1H), 6.86 (t, J_1 = 2.0 Hz, 1H), 6.69 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 6.23 (dd, J_1 = 3.5 Hz, J_2 = 1.7 Hz, 1H), 6.08 (t, J_1 = 2.9 Hz, 1H), 5.27 (s, 1H), 4.96 (s, 1H), 4.41 (s, 1H, $H_{3'}$), 3.73-3.68 (m, 3H, $H_{4'}$ and $2H_{5'}$), 3.67 (s, 3H), 2.71-2.63 (m, 1H, $H_{2'}$), 1.96 (ddd, J_1 = 13.4 Hz, J_2 = 7.4 Hz, J_3 = 3.8 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.5, 149.9, 137.7, 133.7, 131.6, 127.9, 125.6, 124.8, 117.5, 116.6, 108.8, 107.6, 86.5, 83.4, 69.6, 60.9, 36.2, 34.9; **IR** (neat) ν: 3370, 3260, 3193, 3078, 2958, 2921, 1681, 1460, 1328, 1084, 1051, 1005, 989, 937, 908, 816, 707, 681, 538 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 286 (1.9 x 10^4); **HR-ESI MS** (*m/z*): $[M + Na]^+$ calcd for $C_{18}H_{19}N_3O_5Na$ 380.12169, found 380.12139.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(2-furan)-quinazoline-2,4-(3*H*)-dione (50g)

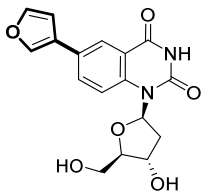


50g

According to the general procedure C: Emissive nucleoside **50g** (0.046 g, 82%) was obtained as a white solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5 → 85:15) using biaryl nucleoside **49g** (0.100 g, 0.163 mmol) and TBAF (0.170 g, 0.650 mmol). The reaction mixture turned purple with the addition of TBAF and a precipitate was observed.

Characterization: R_f (CH_2Cl_2/MeOH, 95:5): 0.11; $[\alpha]_D^{27}$ +29.8 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.70 (s, br, 1H), 8.24 (d, J_1 = 1.8 Hz, 1H), 7.97-7.92 (m, 2H), 7.77 (d, J_1 = 1.2 Hz, 1H), 7.05 (d, J_1 = 3.3 Hz, 1H), 6.69 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 6.61 (dd, J_1 = 3.3 Hz, J_2 = 1.8 Hz, 1H), 5.26 (d, J_1 = 5.0 Hz, 1H), 4.98 (t, J_1 = 5.2 Hz, 1H), 4.42-4.39 (m, 1H, $H_{3'}$), 3.73-3.64 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.68-2.60 (m, 1H, $H_{2'}$), 1.94 (ddd, J_1 = 13.2 Hz, J_2 = 7.3 Hz, J_3 = 3.5 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.3, 151.5, 149.8, 143.2, 138.4, 129.3, 125.3, 121.4, 117.9, 116.9, 112.2, 106.2, 86.5, 83.4, 69.5, 60.8, 36.3; **IR** (neat) ν: 3583, 3483, 3416, 3163, 3036, 2848, 1687, 1662, 1625, 1508, 1487, 1391, 1324, 1091, 1045, 1020, 777, 517 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 290 (3.0×10^4); **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{17}H_{16}N_2O_6Na$ 367.09006, found 367.08983.

1'-(2'-deoxy-β-D-ribofuranoside)-6-(3-furan)-quinazoline-2,4-(3H)-dione (50h)

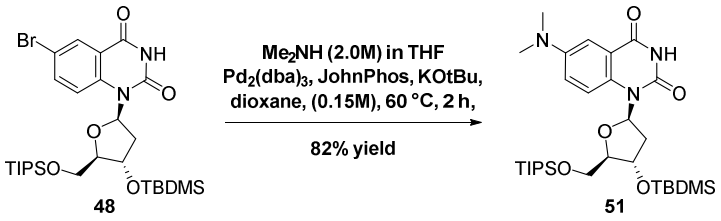


50h

According to the general procedure C: Emissive nucleoside **50h** (0.0402 g, 72%) was obtained as a white solid after column chromatography on silica gel (CH_2Cl_2/MeOH, 95:5 → 90:10) using biaryl nucleoside **49h** (0.100 g, 0.163 mmol) and TBAF (0.170 g, 0.650 mmol). The reaction mixture turned purple with the addition of TBAF.

Characterization: R_f (CH_2Cl_2/MeOH, 95:5): 0.19; $[\alpha]_D^{28}$ +47.8 (*c* 0.05, DMSO); **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.65 (s, br, 1H), 8.30 (s, 1H), 8.17 (s, 1H), 7.89 (d, J_1 = 1.1 Hz, 2H), 7.75 (t, J_1 = 1.7 Hz, 1H), 7.03 (dd, J_1 = 1.7 Hz, J_2 = 0.8 Hz, 1H), 6.69 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.26 (d, J_1 = 5.1 Hz, 1H), 4.98 (t, J_1 = 5.2 Hz, 1H), 4.42-4.39 (m, 1H, $H_{3'}$), 3.73-3.63 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.68-2.60 (m, 1H, $H_{2'}$), 1.94 (ddd, J_1 = 13.6 Hz, J_2 = 7.4 Hz, J_3 = 3.6 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.5, 149.9, 144.5, 139.7, 138.1, 131.6, 127.0, 124.3, 123.5, 117.8, 116.9, 108.5, 86.4, 83.5, 69.5, 60.7, 36.2; **IR** (neat) ν: 3489, 3422, 3188, 3077, 1682, 1518, 1386, 1301, 1041, 803, 591 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 267 (1.9×10^4); **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{17}H_{16}N_2O_6Na$ 367.09061, found 367.09042.

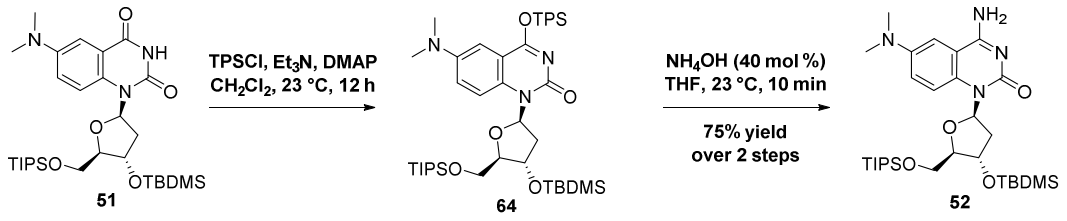
1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-6-(*N,N*-dimethylamino)-quinazoline-2,4-(3*H*)-dione (51)



Procedure: Dioxane (freshly degassed, 0.2 M, 8.0 ml) and Me_2NH (2 M solution in THF, 4.90 mL, 9.56 mmol) were added to a Schlenk tube containing nucleoside **48** (1.00 g, 1.59 mmol), $Pd_2(dba)_3$ (0.073 g, 0.079 mmol), (biphenyl-2-yl)-di-*tert*-butylphosphane (JohnPhos, 0.095 g, 0.319 mmol) and potassium *tert*-butoxide (KOtBu, 0.393 g, 3.505 mmol). The reaction mixture was stirred at 60 °C for 2h. It was then filtered over Celite, washed with CH_2Cl_2 and evaporated *in vacuo*. The residue was subjected to column chromatography on silica gel (Hexane/EtOAc, 9:1 → 7:3) to give **51** (0.772 g, 82%) as a yellow foam.

Characterization: R_f (Hexane/EtOAc, 8:2): 0.19; **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 9.04 (s, br, 1H), 7.66 (d, J_1 = 9.3 Hz, 1H), 7.42 (d, J_1 = 2.9 Hz, 1H), 7.01 (dd, J_1 = 9.3 Hz, J_2 = 2.9 Hz, 1H), 6.75 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 4.73 (quint, J_1 = 4.6 Hz, 1H, $H_{3'}$), 4.04 (dd, J_1 = 11.4 Hz, J_2 = 4.6 Hz, 1H, $H_{5'}$), 3.93 (dd, J_1 = 11.3 Hz, J_2 = 3.1 Hz, 1H, $H_{5'}$), 3.83-3.82 (m, 1H, $H_{4'}$), 2.98 (s, 6H), 2.80-2.73 (m, 1H, $H_{2'}$), 2.05-1.99 (m, 1H, $H_{2'}$), 1.15-1.07 (m, 21H), 0.89 (s, 9H), 0.12 (s, 3H), 0.09 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 162.5, 150.0, 147.0, 130.5, 120.1, 118.2, 117.7, 109.8, 86.9, 84.0, 70.9, 62.0, 40.8, 37.1, 25.9, 18.3, 18.2, 12.2, -5.0, -4.7; **IR** (neat) ν: 3182, 3064, 2941, 2864, 1702, 1685, 1514, 1472, 1378, 1320, 1254, 1110, 1082, 1064, 836, 776 cm^{-1}; **HR-ESI MS** (m/z): $[M + Na]^+$ calcd for $C_{30}H_{53}N_3O_5Si_2Na$ 614.34086, found 614.34092.

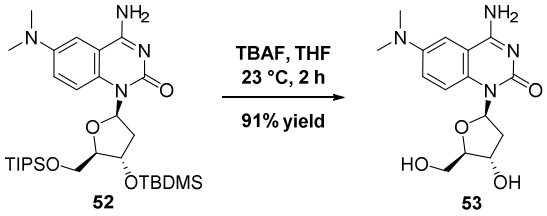
1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-4-amino-6-(*N,N*-dimethylamino)-quinazolone (52)



Procedure: To a stirring solution of **51** (3.00 g, 5.07 mmol) in CH_2Cl_2 (100 mL), was added Et_3N (1.40 mL, 10.14 mmol). Subsequently, 2,4,6-triisopropylbenzenesulfonyl chloride (TPSCl, 3.07 g, 10.14 mmol) and DMAP (1.24 g, 10.14 mmol) were added and the reaction was stirred at room temperature for 12h. The deep red reaction solution was then quenched with sat. sol. NH_4Cl and extracted with EtOAc (3 times). The organic layer was dried over $MgSO_4$, filtered and evaporated *in vacuo* to give **64** as a red foam. This material was used in the next step without further purification. Nucleoside **64** was then dissolved in THF (60 mL). A solution of NH_4OH (30 mL) was added and the reaction was stirred at room temperature for 10 min. The yellow reaction was then quenched by the addition of NH_4Cl and was extracted with EtOAc (3 times). The combined organic layer was washed with $NaHCO_3$ (3 times), dried over $MgSO_4$, filtered and evaporated *in vacuo*. In order to remove residual 2,4,6-triisopropyl-benzenesulfonic acid, the residue was dissolved in MeOH and ion-exchange resin (Dowex 1X4 chloride form) was added and the solution was stirred 12h at room temperature. It then was filtered, evaporated *in vacuo* and subjected to column chromatography on silica gel (EtOAc/MeOH, 100:0 → 92:8) to give **52** (2.25 g, 75% over two steps) as a yellow foam.

Characterization: $\mathbf{R_f}$ (EtOAc /MeOH, 96:4): 0.35; **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 7.70 (d, J_1 = 9.4 Hz, 1H), 7.00 (dd, J_1 = 9.4 Hz, J_2 = 2.7 Hz, 1H), 6.93 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 6.91 (s, br, 1H), 4.73-4.68 (m, 1H, $H_{3'}$), 4.04 (dd, J_1 = 11.3 Hz, J_2 = 2.8 Hz, 1H, $H_{5'}$), 3.92 (dd, J_1 = 11.3 Hz, J_2 = 3.2 Hz, 1H, $H_{5'}$), 3.79-3.77 (m, 1H, $H_{4'}$), 2.89 (s, 6H), 2.73-2.66 (m, 1H, $H_{2'}$), 1.99 (ddd, J_1 = 13.3 Hz, J_2 = 7.5 Hz, J_3 = 3.6 Hz, 1H, $H_{2'}$), 1.16-1.07 (m, 21H), 0.87 (s, 9H), 0.07 (s, 3H), 0.05 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 163.5, 156.3, 146.2, 132.8, 120.1, 118.0, 112.0, 105.5, 86.07, 84.3, 70.4, 62.1, 40.9, 37.2, 25.9, 18.2, 18.1, 12.1, -4.4, -4.7; **IR** (neat) ν: 3285, 3089, 2927, 2894, 2864, 1613, 1539, 1464, 1332, 1256, 1142, 1106, 1061, 1034, 881, 835, 775, 677 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + H]^+$ calcd for $C_{30}H_{55}N_4O_4Si_2$ 591.37564, found 591.37552.

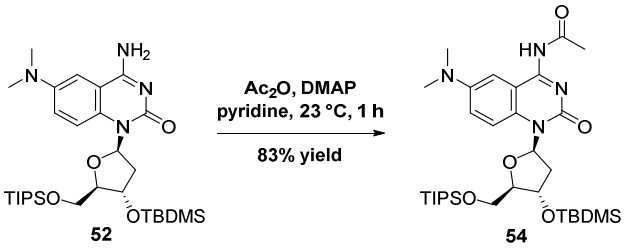
1'-(2'-deoxy-β-D-ribofuranoside)-4-amino-6-(*N,N*-dimethylamino)-quinazolone (53)



Procedure: To a stirred solution of the nucleoside **52** (0.080 g, 0.136 mmol) in THF (2 mL), was added TBAF·3H_2O (0.142 g, 0.542 mmol) dissolved in THF. The mixture was stirred at room temperature for 2h and evaporated *in vacuo*. The residue was purified by column chromatography on silica gel (MeCN/H_2O, 98:2 → 92:8). The residue was then washed with H_2O (3x) and finally dried under high vacuum to obtain nucleoside **53** (0.039 g, 91%) as a yellow solid.

Characterization: R_f (MeCN/H_2O, 92:8): 0.28; **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 9.55 (s, br, 1H), 8.80 (s, br, 1H), 7.78 (d, J_1 = 9.5 Hz, 1H), 7.45 (d, J_1 = 2.2 Hz, 1H), 7.20 (dd, J_1 = 9.4 Hz, J_2 = 2.3 Hz, 1H), 6.67 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 5.29 (s, br, 1H), 4.99 (s, br, 1H), 4.38 (s, 1H, $H_{3'}$), 3.71-3.61 (m, 3H, $H_{4'}$ and 2$H_{5'}$), 2.94 (s, 6H), 2.62-2.54 (m, 1H, $H_{2'}$), 1.92-1.86 (m, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 160.7, 150.8, 146.2, 131.2, 121.0, 118.0, 110.3, 106.7, 86.5, 83.9, 69.7, 60.9, 40.5, 36.4; **IR** (neat) ν: 3333, 3189, 2899, 2677, 1715, 1667, 1574, 1537, 1347, 1054, 1015, 991, 797, 604 cm^{-1}; **UV-Vis** (H_2O): λ_{abs} (nm) and ε (M^{-1} cm^{-1}) = 260 and 17000, 365 and 2500; **HR-ESI MS** (*m/z*): $[M + H]^+$ calcd for $C_{15}H_{21}N_4O_4$ 321.15573, found 321.15558.

1'-(3'-*O*-tertbutyldimethylsilyl-5'-*O*-triisopropylsilyl-2'-deoxy-β-D-ribofuranoside)-4-(*N*-acetyl)-amino-6-(*N,N*-dimethylamino)-quinazolone (54)

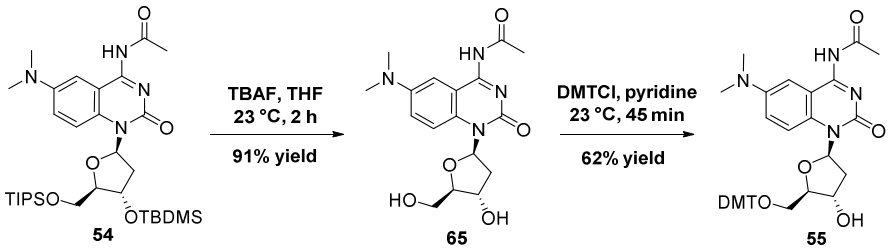


Procedure: To a stirring solution of **52** (1.83 g, 3.10 mmol) in pyridine (40 mL), was added DMAP (0.038 g, 0.31 mmol) and Ac_2O (0.38 mL, 4.03 mmol). The reaction mixture was stirred at room

temperature for 1h and then quenched with MeOH. The solution was diluted with EtOAc and washed with HCl (0.05 M) and a sat sol. $CuSO_4$. The combined aqueous layer were back extracted with EtOAc and the combined organic layer was washed with a sat sol. NaCl, dried over $MgSO_4$, filtered, evaporated *in vacuo* and co-evaporated with toluene. The residue was then subjected to column chromatography on silica gel (hexane/EtOAc, 9:1 → 8:2) to give **54** (1.64 g, 83%) as a yellow foam.

Characterization: R_f (hexane/EtOAc, 8:2): 0.24; **^{1}H NMR** (400 MHz, $CDCl_3$) δ: 12.77 (s, br, 1H), 7.66-7.64 (m, 2H), 7.05 (dd, J_1 = 9.4 Hz, J_2 = 3.2 Hz, 1H), 6.71 (t, J_1 = 7.7 Hz, 1H, $H_{1'}$), 4.75-4.70 (m, 1H, $H_{3'}$), 4.04 (dd, J_1 = 11.3 Hz, J_2 = 3.0 Hz, 1H, $H_{5'}$), 3.92 (dd, J_1 = 11.3 Hz, J_2 = 3.4 Hz, 1H, $H_{5'}$), 3.84-3.81 (m, 1H, $H_{4'}$), 2.99 (s, 6H), 2.82-2.74 (m, 1H, $H_{2'}$), 2.35 (s, 3H), 1.03 (dd, J_1 = 13.3 Hz, J_2 = 7.5 Hz, J_3 = 3.8 Hz, 1H, $H_{2'}$), 1.16-1.08 (m, 21H), 0.89 (s, 9H), 0.09 (s, 3H), 0.07 (s, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ: 188.0, 157.1, 147.0, 131.3, 120.9, 117.9, 116.9, 108.6, 86.6, 84.3, 70.4, 62.1, 40.8, 37.2, 28.0, 25.9, 18.2, 18.1, 12.2, -4.4, -4.6; **IR** (neat) ν: 2941, 2895, 2864, 1698, 1616, 1586, 1562, 1508, 1277, 1252, 1107, 1063, 997, 881, 835, 777 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{32}H_{56}N_4O_5Si_2Na$ 655.36815, found 655.36829.

1'-[5'-*O*-(4,4-dimethoxytrityl)-2'-deoxy-β-D-ribofuranoside]-4-(*N*-acetyl)-amino-6-(*N*,*N*-dimethylamino)-quinazolone (55)



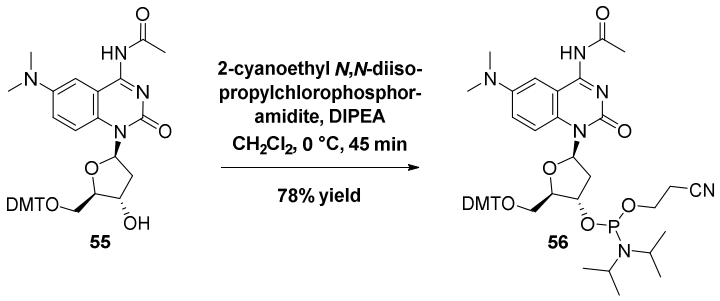
54 65 55

Procedure: To a stirred solution of the nucleoside **54** (0.570 g, 0.90 mmol) in THF (14 mL), was added TBAF·3H_2O (0.942 g, 3.60 mmol) dissolved in THF. The mixture was stirred at room temperature for 2h, quenched with MeOH, and evaporated *in vacuo*. The residue was purified by column chromatography on silica gel (MeCN/H_2O, 100:0 → 98:2) to obtain nucleoside **65** (0.297 g, 91%) as an orange foam. R_f (MeCN/H_2O, 98:2): 0.24. The material was directly used in the next step without further purification. Nucleoside **65** (0.168 g, 0.464 mmol) was co-evaporated with pyridine (2 x 2 mL) and then

dissolved in pyridine (2.5 mL). To the stirring solution was added dropwise DMTCl (0.188 g, 0.556 mmol) dissolved in pyridine. The reaction mixture was stirred at room temperature for 45 min and quenched with a sat. sol. $NaHCO_3$. The resulting mixture was extracted with CH_2Cl_2 and the combined organic layer was dried over $MgSO_4$, filtered, evaporated *in vacuo* and co-evaporated with toluene. The residue was purified by column chromatography on silica gel (CH_2Cl_2/MeOH/Et_3N, 99:0.5:0.5 → 98.5:1:0.5) to obtain nucleoside **55** (0.190 g, 62%, containing 3% of Et_3NH adduct, 80% brsm) as a yellow-orange foam along with the recovered starting material **65** (0.039, 12%).

Characterization: R_f (CH_2Cl_2/MeOH/Et_3N, 95:5:0.5): 0.32; **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 10.32 (s, br, 1H), 7.86 (d, J_1 = 9.5 Hz, 1H), 7.40-7.37 (m, 2H), 7.32 (d, J_1 = 2.9 Hz, 1H), 7.30-7.21 (m, 7H), 6.84 (dd, J_1 = 8.9 Hz, J_2 = 7.8 Hz, 4H), 6.71 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 6.55 (dd, J_1 = 9.4 Hz, J_2 = 3.0 Hz, 1H), 5.34 (dd, J_1 = 5.2 Hz, 1H), 4.64-4.58 (m, 1H, $H_{3'}$), 3.88-3.85 (m, 1H, $H_{4'}$), 3.71 (s, 6H), 3.37-3.28 (m, 2H, $H_{5'}$, expected), 2.81 (s, 6H), 2.76-2.69 (m, 1H, $H_{2'}$), 2.35 (s, 3H), 2.02-1.96 (m, 1H, $H_{2'}$); **^{13}C NMR** (125 MHz, DMSO-d_6) δ: 158.09, 158.06, 145.9, 144.5, 135.5, 135.3, 131.6, 129.84, 129.81, 127.9, 127.8, 126.8, 120.5, 117.6, 113.1, 106.7, 85.6, 84.8, 84.0, 69.7, 62.6, 54.97, 54.96, 45.7, 40.1, 36.7, 26.8, 11.6; **IR** (neat) ν: 3417, 2951, 2932, 2841, 1692, 1607, 1583, 1561, 1507, 1279, 1249, 1175, 1032, 827, 738 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{38}H_{40}N_4O_7Na$ 687.27892, found 687.27812.

1'-[3'-*O*-[2-cyanoethoxy-(*N*,*N*-diisopropylamino)-phosphino]-(5'-*O*-(4,4-dimethoxytrityl)-2'-deoxy-β-D-ribofuranoside]-4-(*N*-acetyl)-amino-6-(*N*,*N*-dimethylamino)-quinazolone (56)

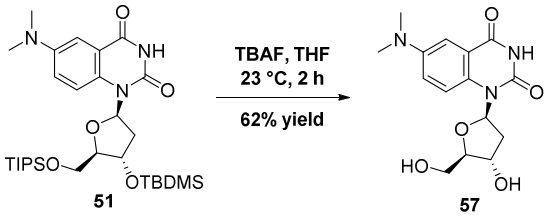


Procedure: To a stirred solution of nucleoside **55** (0.100 g, 0.150 mmol) in CH_2Cl_2 (1.0 mL) at 0 °C, was added freshly distilled diisopropylethylamine (DIPEA, 65 µL, 0.376 mmol) and the reaction was stirred 5 min at 0 °C. To this solution was added 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (67 µL, 0.301 mmol) and the reaction was stirred at room temperature for 45 min. The reaction was

quenched with a sat. sol. $NaHCO_3$ and extracted with CH_2Cl_2. The combined organic layer was dried over $MgSO_4$, filtered, evaporated *in vacuo*. The residue was purified by column chromatography on silica gel (Hexane/EtOAc/Et_3N, 60:40:0.5 → 55:45:0.5) to obtain nucleoside **56** (0.100 g, 78%, diastereomeric mixture) as a yellow foam.

Characterization: R_f (Hexane/EtOAc/Et_3N, 40:60:0.5): 0.38 and 0.47; **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 7.87-7.84 (m, 2H), 7.41-7.37 (m, 4H), 7.34 (s, 2H), 7.28-7.21 (m, 16H), 6.84-6.78 (m, 8H), 6.71-6.67 (m, 2H, $H_{1'}$), 6.57-6.61 (m, 2H), 4.97-4.83 (m, 2H, $H_{3'}$), 4.08-4.02 (m, 2H, $H_{4'}$), 3.76-3.71 (m, 2H), 3.70 (s, 12H), 3.62-3.47 (m, 4H), 3.37-3.35 (m, 2H, $H_{5'}$), 2.89-2.84 (m, 2H, $H_{2'}$), 2.82 (s, 12H), 2.74 (t, J_1 = 5.9 Hz, 2H), 2.62 (t, J_1 = 5.9 Hz, 2H), 2.35 (br, s, 6H), 2.29-2.22 (m, 2H, $H_{2'}$), 1.34-1.25 (m, 4H), 1.12 (d, J_1 = 6.8 Hz, 6H), 1.10-1.08 (m, 12H), 0.95 (d, J_1 = 6.8 Hz, 6H); **^{13}C NMR** (125 MHz, DMSO-d_6) δ: 170.3, 166.9, 161.8, 158.13, 158.12, 158.09, 149.59, 149.56, 146.2, 146.0, 144.43, 144.41, 135.4, 135.3, 135.2, 135.16, 131.7, 131.5, 129.81, 129.77, 128.6, 128.0, 127.9, 127.74, 127.73, 126.79, 126.77, 120.6, 120.5, 118.9, 118.7, 117.24, 117.19, 113.06, 113.05, 106.7, 85.7, 85.68, 84.2, 84.1, 84.0, 83.9, 83.8, 83.7, 72.6, 72.4, 71.8, 71.7, 67.4, 62.4, 62.2, 59.7, 58.44, 58.37, 58.3, 58.2, 54.95, 54.94, 54.93, 54.87, 42.6, 42.57, 42.5, 42.47, 40.1, 38.1, 36.0, 35.7, 29.8, 28.3, 24.3, 24.27, 24.26, 24.23, 24.22, 24.18, 24.16, 24.1, 23.2, 22.4, 20.7, 19.83, 19.78, 19.76, 19.7, 14.0, 13.8, 10.8; **^{31}P NMR** (168 MHz, DMSO-d_6) δ: 147.9 , 147.3; **IR** (neat) ν: 2961, 2932, 2874, 1697, 1610, 1585, 1562, 1508, 1280, 1251, 1178, 1034, 828 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{47}H_{57}N_6O_8PNa$ 887.38677, found 887.38622.

1'-(2'-deoxy-β-d-ribofuranoside)-6-(*N*,*N*-dimethylamino)-quinazoline-2,4-(3*H*)-dione (57)

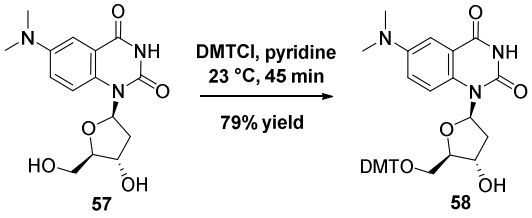


Procedure: To a stirred solution of the nucleoside **51** (1.50 g, 2.53 mmol) in THF (36 mL), was added TBAF (2.65 g, 10.14 mmol) dissolved in THF. The mixture was stirred at room temperature for 2h and evaporated *in vacuo*. The residue was then washed with MeOH (6.0 mL), aqueous solution of

HOAc (6.0 mL) and H_2O (6.0 mL) and finally dried under high vacuum to obtain nucleoside **57** (0.500, 62%) as a yellow solid.

Characterization: R_f (CH_2Cl_2/MeOH, 94:6): 0.23; **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.49 (s, br, 1H), 7.72 (d, J_1 = 9.3 Hz, 1H), 7.19 (d, J_1 = 3.2 Hz, 1H), 7.11 (dd, J_1 = 9.3 Hz, J_2 = 3.2 Hz, 1H), 6.64 (t, J_1 = 7.8 Hz, 1H, $H_{1'}$), 5.23 (d, J_1 = 5.0 Hz, 1H), 4.93 (t, J_1 = 5.0 Hz, 1H), 4.38-4.35 (m, 1H, $H_{3'}$), 3.71-3.61 (m, 3H, $H_{4'}$ and $2H_{5'}$), 2.92 (s, 6H), 2.63-2.56 (m, 1H, $H_{2'}$), 1.89 (ddd, J_1 = 13.3 Hz, J_2 = 7.3 Hz, J_3 = 3.5 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.8, 149.7, 146.9, 129.8, 119.7, 118.2, 117.1, 108.1, 86.3, 83.2, 69.6, 60.9, 40.5, 36.1; **IR** (neat) ν: 3474, 3411, 3197, 3067, 2885, 2812, 1668, 1581, 1515, 1483, 1387, 1387, 1308, 1069, 1051, 1040, 1014, 1005, 810, 769, 626, 540, 522 cm^{-1}; **UV-Vis** (H_2O): λ_{max} (nm) and ε (M^{-1} cm^{-1}) = 356 and ; **HR-ESI MS** (m/z): $[M + Na]^+$ calcd for $C_{15}H_{19}N_3O_5Na$ 344.12169, found 344.12193.

1'-[5'-*O*-(4,4-dimethoxytrityl)-2'-deoxy-β-d-ribofuranoside]-6-(*N*,*N*-dimethylamino)-quinazoline-2,4-(3*H*)-dione (58)



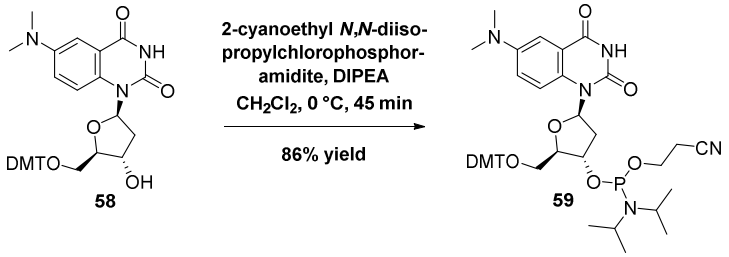
57 58

Procedure: Nucleoside **57** (0.150 g, 0.467 mmol) was co-evaporated with pyridine (2 x 1.5 mL) and then suspended in pyridine (2.0 mL). To the stirring solution was added dropwise DMTCl (0.190 g, 0.560 mmol) dissolved in pyridine. A clear red reaction mixture was observed and it was stirred at room temperature for 45 min. The reaction was then quenched with a sat. sol. $NaHCO_3$. The resulting mixture was extracted with CH_2Cl_2 and the combined organic layer was dried over $MgSO_4$, filtered, evaporated *in vacuo* and co-evaporated with toluene. The residue was purified by column chromatography on silica gel (CH_2Cl_2/MeOH/Et_3N, 99:0.5:0.5 → 98.5:1:0.5) to obtain nucleoside **58** (0.229 g, 79%, containing 1% of Et_3N) as a green-yellow solid.

Characterization: R_f (CH_2Cl_2/MeOH/Et_3N, 95:5:0.5): 0.39; **^{1}H NMR** (400 MHz, DMSO-d_6) δ: 11.48 (s, br, 1H), 7.81 (d, J_1 = 9.3 Hz, 1H), 7.42 (d, J_1 = 7.2 Hz, 2H), 7.30-7.27 (m, 6H), 7.24-7.19 (m,

2H), 6.87-6.83 (m, 4H), 6.71 (t, J_1 = 7.6 Hz, 1H, $H_{1'}$), 6.48 (dd, J_1 = 9.4 Hz, J_2 = 3.2 Hz, 1H), 5.38 (s, br, 1H), 4.65-4.63 (m, 1H, $H_{3'}$), 3.88-3.86 (m, 1H, $H_{4'}$), 3.70 (s, 6H), 3.39-3.31 (m, 2H, $2H_{5'}$), 2.79 (s, 6H), 2.76-2.71 (m, 1H, $H_{2'}$), 1.98 (ddd, J_1 = 13.4 Hz, J_2 = 7.2 Hz, J_3 = 3.4 Hz, 1H, $H_{2'}$); **^{13}C NMR** (100 MHz, DMSO-d_6) δ: 161.9, 158.11, 158.09, 149.7, 146.2, 144.6, 135.5, 135.4, 129.85, 129.83, 129.79, 127.96, 127.8, 126.8, 119.5, 118.0, 117.2, 113.1, 108.2, 85.6, 84.6, 83.3, 69.6, 62.6, 55.0, 45.6, 40.0, 36.4, 10.7; **IR** (neat) ν: 3412, 3183, 3055, 2940, 1682, 1508, 1481, 1316, 1249, 1176, 1076, 1032, 827, 731 cm^{-1}; **HR-ESI MS** (*m*/*z*): $[M + Na]^+$ calcd for $C_{36}H_{37}N_3O_7Na$ 646.25237, found 646.25220.

1'-[3'-*O*-[2-cyanoethoxy-(*N*,*N*-diisopropylamino)-phosphino]-(5'-*O*-(4,4-dimethoxytrityl)-2'-deoxy-β-d-ribofuranoside]-6-(*N*,*N*-dimethylamino)-quinazoline-2,4-(3*H*)-dione (59)



Procedure: To a stirred solution of nucleoside **58** (0.130 g, 0.208 mmol) in CH_2Cl_2 (1.2 mL) at 0 °C, was added freshly distilled diisopropylethylamine (DIPEA, 91 µL, 0.521 mmol) and the reaction was stirred 5 min at 0 °C. To this solution was added 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (93 µL, 0.416 mmol) and the reaction was stirred at room temperature for 45 min. The reaction was quenched with a sat. sol. $NaHCO_3$ and extracted with CH_2Cl_2. The combined organic layer was dried over $MgSO_4$, filtered, evaporated *in vacuo*. The residue was purified by column chromatography on silica gel (Hexane/EtOAc/Et_3N, 60:40:0.5 → 55:45:0.5) to obtain nucleoside **59** (0.147 g, 86%, diastereomeric mixture) as a yellow foam.

Characterization: **R_f** (Hexane/EtOAc/Et_3N, 40:60:0.5): 0.32 and 0.42; **^{1}H NMR** (500 MHz, DMSO-d_6) δ: 11.48 (s, br, 2H), 7.81-7.78 (m, 2H), 7.42-7.38 (m, 4H), 7.29-7.20 (m, 16H), 6.85-6.79 (m, 8H), 6.71-6.67 (m, 2H, $H_{1'}$), 6.57-6.53 (m, 2H), 4.92-4.86 (m, 2H, $H_{3'}$), 4.05-4.00 (m, 2H, $H_{4'}$), 3.76-3.71 (m, 2H), 3.70 (s, 12H), 3.59-3.47 (m, 4H), 3.37-3.35 (m, 2H, $H_{5'}$), 2.89-2.84 (m, 2H, $H_{2'}$), 2.80 (s, 12H), 2.74 (t, J_1 = 5.9 Hz, 2H), 2.62 (t, J_1 = 5.9 Hz, 2H), 2.26-2.15 (m, 2H, $H_{2'}$), 1.34-1.25 (m, 4H), 1.12 (d, J_1 =

6.8 Hz, 6H), 1.10-1.08 (m, 12H), 0.95 (d, J_1 = 6.8 Hz, 6H); **^{13}C NMR** (125 MHz, DMSO-d_6) δ: 170.2, 166.9, 161.85, 161.84, 158.14, 158.13, 158.11, 149.60, 149.57, 146.2, 144.45, 144.41, 135.4, 135.3, 135.2, 135.1, 131.7, 131.5, 129.96, 129.95, 129.81, 129.79, 128.6, 127.94, 127.91, 127.75, 127.74, 126.78, 126.77, 119.6, 119.5, 118.9, 118.7, 117.67, 117.65, 117.2, 117.1, 113.07, 113.06, 108.2, 85.72, 85.71, 83.69, 83.66, 83.50, 83.39, 72.4, 72.3, 71.6, 71.5, 67.3, 62.3, 62.1, 59.7, 58.45, 58.38, 58.31, 58.23, 54.95, 54.94, 54.93, 54.86, 42.59, 42.58, 42.49, 42.47, 40.0, 38.1, 35.8, 35.5, 29.8, 28.3, 24.3, 24.23, 24.21, 24.18, 24.16, 24.10, 23.2, 22.4, 20.7, 19.8, 19.78, 19.77, 19,71, 14.0, 13.8, 10.8; **^{31}P NMR** (202 MHz, DMSO-d_6) δ: 147.8, 147.3; **IR** (neat) ν: 3178, 3069, 2964, 2930, 2873, 1698, 1684, 1509, 1379, 1319, 1251, 1178, 1073, 1031, 977, 829 cm^{-1}; **HR-ESI MS** (m/z): $[M + Na]^+$ calcd for $C_{45}H_{54}N_5O_8PNa$ 846.36022, found 846.36021.

7.2. Photophysical Properties of the Nucleosides

7.2.1. General Measurements Details

For all photophysical experiments, the fluorescent nucleosides were prepared as 20 mM DMSO stock solutions and stored at -20 °C. All measurements were collected using a Molecular Devices Spectra spectrophotometer in a quartz cuvette (l = 1 cm) and were performed in duplicate at a concentration of 40 μM of nucleoside. Measurements were performed with spectrophotometric grade dioxane and deionized H_2O. For all values reported, final DMSO concentrations after dilution were always less than 0.2%. E_T^{30} values for dioxane / water mixtures were determined by dissolving a small amount of Reichardt's dye in the solution and measuring the most red-shifted absorption maximum (**Table 7.1**).[219] The value was converted into E_T^{30} using the formula: $E_T^{30} = 28591/\lambda_{abs}$.

Molar extinction coefficient measurements were conducted in deionized H_2O using Beer's Law plots of concentration versus absorbance. Quantum yield measurements were performed in triplicate at the most red-shifted absorbance wavelength of each nucleoside and corresponding standard using optical densities of 0.1 ± 0.05. 2-Aminopyridine (ϕ = 0.60 at λ_{ex} = 285 nm) in 0.5 M H_2SO_4 (n = 1.432) was used as the fluorescent standard for the relative quantum yields (ϕ) of **50a** – **h**.[215] Quinine hemisulfate (ϕ = 0.546 at λ_{ex} = 370 nm) in 0.5 M H_2SO_4 (n = 1.432) was used as the fluorescent standard for the relative quantum yields of **^{DMA}C**.[259] 2-Aminopyridine was used to cross-referenced the standard to give ϕ = 0.59,

which is in good agreement with the literature reported value.[215] Quantum yields were calculated as described, according to the following equation (eq. 1):

$$\phi = \phi_R \frac{F}{F_R} \frac{A_R}{A} \frac{n^2}{n_R^2}$$

where φ_R is the quantum yield of the reference, F and F_R are the integrated emission intensities of the probe and the reference, respectively. A and A_R are the optical densities of the probe and the reference, respectively. n and n_R are the refractive indexes of the solvent in which the measurement was performed for the probe and the reference, respectively.

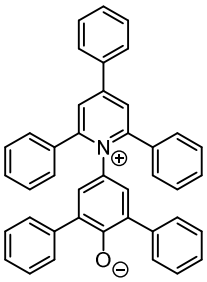
Reichardt's dye	Solvent	λ_{abs} [a]	λ_{abs} [b]	E_T^{30} [c]	E_T^{30} [d]	E_T^{30} [e]
	Dioxane	788	782	36.28	36.56	36.42
	0.90 - 0.10	615	642	46.49	44.53	45.51
	0.70 - 0.30	562	562	50.87	50.87	50.87
	0.30 - 0.70	500	500	57.18	57.18	57.18
	Water	453	453	63.11	63.11	63.11

Table 7.1. Absorbance maxima of Reichardt's dye in solvent mixtures of dioxane and water. The value in pure water was obtained from literature because of the limited solubility of the dye in water.[219] λ_{abs} are reported at the most red-shifted absorbance wavelength in nm from [a] first reading and [b] second reading. E_T^{30} are reported in kcal mol^{-1} and calculated from [c] first reading and [d] second reading. [e] The final E_T^{30} value was averaged from the two independent measurements.

7.3. Molecular Modeling

7.3.1. Calculations for Biaryl Nucleosides 50a – h

Electronic properties were calculated from DFT optimized geometries of the isolated nucleobases of nucleosides **50a – h**. The global geometric minimum for each nucleobase was used to calculate the molecular orbital plots and energies using a low energy level of calculation, with values corresponding to a ground state (S_0) assigned to 0 kJ mol^{-1}. Calculations were performed using Spartan 10.

7.3.2. Calculations for ^{DMA}C and ^{DMA}T

The optimized geometries for the *syn* and *anti* conformations of ^{DMA}**C** (**53**) and ^{DMA}**T** (**57**) were calculated as followed. An equilibrium geometry calculation using Hartree-Fock at the 3-21G level was performed to obtain 81 conformations of ^{DMA}**C**. Then, the minimum geometries corresponding to the *syn* and *anti* conformation of ^{DMA}**C** and ^{DMA}**T** were further optimized using DFT and B3LYP/6-311++G**. Electronic properties and molecular orbitals plots were calculated from DFT optimized geometries of the isolated nucleobase of ^{DMA}**C** using B3LYP/6-311++G**. All calculations were performed using Spartan 10.

7.4. Oligonucleotides

7.4.1. Synthesis and Purification

The wild-type sequences were obtained from Sigma-Aldrich as HPLC purified products. ^{DMA}**C**-modified oligonucleotides were synthesized on a 1 μmol scale using a DNA synthesizer according to the standard Trityl-off procedure, except that 5% dichloroacetic acid in CH_2Cl_2 was used for Trityl deprotection. Two coupling reactions were performed for the site-specific introduction of the modified nucleoside into oligonucleotides. The phosphoramidite **56** (0.160 g, 0.185 mmol) was dissolved in dry acetonitrile (1.6 mL) and was added to the DNA synthesizer. The synthesis of the oligonucleotide was monitored using DMT deprotection. Upon completion of the sequences, the oligonucleotides were cleaved from the solid support and deprotected by treatment with 1.2 mL of 33% aqueous ammonium hydroxide at 55 °C for 12 hours in a 1.5 mL screw-top cap tube. The resulting products were lyophilized, and HPLC-purified with a semi-prep C-18 reverse-phase column (YMCbasic B-22-10P 150 x 10 mm) using a *Varian Pro Star* HPLC system. The gradient condition was 5% to 10% acetonitrile over 50 minutes at 3 mL/min in 0.1 M triethylammonium acetate (TEAA, pH 7.0). Peaks were collected, lyophilized to dryness and HPLC-analyzed with an analytical column (XBridge Phenyl 3.5 μm 150 x 4.16 mm). The gradient condition was 5% to 18% acetonitrile over 40 minutes at 0.5 mL/min in TEAA. DNA samples were analyzed using HR-ESI mass spectroscopy.

7.4.2. Quantification

Oligonucleotide stock solutions were prepared as a 100 μM concentration in pure water and quantified using their calculated molar extinction coefficient (ε) at 260 nm. The extinction coefficients of the non-modified DNA were calculated using the nearest-neighbour model at 260 nm.[311-312] The following values were used in $cm^{-1} M^{-1}$: ε(**TeloC**) = 185 900, ε(**TeloG**) = 215 000, ε(**TeloT2**) = 186 800, ε(**TeloA2**) = 191 100, ε(**TeloT13**) = 187 800, ε(**TeloA13**) = 190 700, ε(**TeloT14**) = 186 800, ε(**TeloA14**) = 191 100, ε(**TeloT15**) = 186 800, ε(**TeloA15**) = 192 100, ε(**TeloG2**) = 218 100, ε(**TeloG13**) = 216 900, ε(**TeloG14**) = 218 100, ε(**TeloG15**) = 217 500, where the letter and the number of the oligonucleotide correspond to the nucleoside and the position, respectively, modified from the wild-type sequences: (**TeloC**): 5'-$(CCCTAA)_3CCC$-3' and (**TeloG**): 5'-$GGG(TTA\ GGG)_3$-3'. The extinction coefficients of the ^{DMA}**C**- and ^{DMA}**T**-modified oligonucleotides were calculated using the base composition method which sums the coefficients of the isolated nucleosides which is then multiplied by a factor of 0.9 to account for the base stacking in the oligonucleotide. The following value was used for the ^{DMA}**C**- and ^{DMA}**T**-modified oligonucleotides: ε(**Telo^{DMA}C**) = 195 210, ε(**Telo^{DMA}T**) = 193 410 $cm^{-1} M^{-1}$.

7.4.3. Oligonucleotides Sequences

Name	Sequence (5'-3')
TeloC	CCC TAA CCC TAA CCC TAA CCC
TeloX2	C**X**C TAA CCC TAA CCC TAA CCC
TeloT2	C**T**C TAA CCC TAA CCC TAA CCC
TeloA2	C**A**C TAA CCC TAA CCC TAA CCC
TeloX13	CCC TAA CCC TAA **X**CC TAA CCC
TeloT13	CCC TAA CCC TAA **T**CC TAA CCC
TeloA13	CCC TAA CCC TAA **A**CC TAA CCC
TeloG2	G**A**G TTA GGG TTA GGG TTA GGG
TeloG13	GGG TTA GGG TTA **A**GG TTA GGG

Name	Sequence (5'-3')
TeloG	GGG TTA GGG TTA GGG TTA GGG
TeloX14	CCC TAA CCC TAA C**X**C TAA CCC
TeloT14	CCC TAA CCC TAA C**T**C TAA CCC
TeloA14	CCC TAA CCC TAA C**A**C TAA CCC
TeloX15	CCC TAA CCC TAA CC**X** TAA CCC
TeloT15	CCC TAA CCC TAA CC**T** TAA CCC
TeloA15	CCC TAA CCC TAA CC**A** TAA CCC
TeloG14	GGG TTA GGG TTA G**A**G TTA GGG
TeloG15	GGG TTA GGG TTA GG**A** TTA GGG

Table 7.2. Names and sequences of DNAs used for thermodynamic experiments (**X** = ^{DMA}**C** or ^{DMA}**T**). **TeloC** and **TeloG** correspond to the human telomeric wild-type sequence and the no. correspond to the site of modification.

Name	Sequence (5'-3')	Molecular Formula	Calculated 10x charged m/z	Measured 10x charged m/z
TeloX2	CXC TAA CCC TAA CCC TAA CCC	$C_{205}H_{257}N_{73}O_{120}P_{20}$	628.01059	628.01135
TeloX13	CCC TAA CCC TAA XCC TAA CCC	$C_{205}H_{257}N_{73}O_{120}P_{20}$	628.01059	628.01020
TeloX14	CCC TAA CCC TAA CXC TAA CCC	$C_{205}H_{257}N_{73}O_{120}P_{20}$	628.01059	628.01092
TeloX15	CCC TAA CCC TAA CCX TAA CCC	$C_{205}H_{257}N_{73}O_{120}P_{20}$	628.01059	628.01160

Name	Sequence (5'-3')	Molecular Formula	Calculated neutral m/z	Measured neutral m/z
TeloX2	CXC TAA CCC TAA CCC TAA CCC	$C_{205}H_{267}N_{73}O_{120}P_{20}$	6290.18	6290.18
TeloX13	CCC TAA CCC TAA XCC TAA CCC	$C_{205}H_{267}N_{73}O_{120}P_{20}$	6290.18	6290.19
TeloX14	CCC TAA CCC TAA CXC TAA CCC	$C_{205}H_{267}N_{73}O_{120}P_{20}$	6290.18	6290.19
TeloX15	CCC TAA CCC TAA CCX TAA CCC	$C_{205}H_{267}N_{73}O_{120}P_{20}$	6290.18	6290.18

Table 7.3. Calculated and measured high resolution HR-ESI MS for the ^{DMA}**C**-containing DNA (**X** = ^{DMA}**C**). (up) Monoisotopic masses obtained for the 10 times negatively charged DNAs. (down) Neutral deconvoluted masses. MS values are reported in m/z.

Name	Sequence (5'-3')	Molecular Formula	Calculated 10x charged m/z	Measured 10x charged m/z
TeloX13	CCC TAA CCC TAA XCC TAA CCC	$C_{205}H_{256}N_{72}O_{121}P_{20}$	628.10899	628.10894
TeloX14	CCC TAA CCC TAA CXC TAA CCC	$C_{205}H_{256}N_{72}O_{121}P_{20}$	628.10899	628.10870
TeloX15	CCC TAA CCC TAA CCX TAA CCC	$C_{205}H_{256}N_{72}O_{121}P_{20}$	628.10899	628.10842

Name	Sequence (5'-3')	Molecular Formula	Calculated neutral m/z	Measured neutral m/z
TeloX13	CCC TAA CCC TAA XCC TAA CCC	$C_{205}H_{266}N_{72}O_{121}P_{20}$	6291.16	6291.17
TeloX14	CCC TAA CCC TAA CXC TAA CCC	$C_{205}H_{266}N_{72}O_{121}P_{20}$	6291.16	6291.17
TeloX15	CCC TAA CCC TAA CCX TAA CCC	$C_{205}H_{266}N_{72}O_{121}P_{20}$	6291.16	6291.17

Table 7.4. Calculated and measured high resolution HR-ESI MS for the ^{DMA}**T**-containing DNA (**X** = ^{DMA}**T**). (up) Monoisotopic masses obtained for the 10 times negatively charged DNAs. (down) Neutral deconvoluted masses. MS values are reported in m/z.

7.4.4. DNA Folding and Buffer Conditions

Oligonucleotide stock solutions were prepared in water at a concentration of 100 μM and were then diluted into the desired buffer to the final concentration. The folding of single-stranded oligonucleotides was controlled by variation of the pH of the buffer. *i*-motif DNA studies were performed in a buffer of 0.2 M Na_2HPO_4, 0.1 M Citric acid and 0.1 M NaCl. Two stock solutions of 0.2 M Na_2HPO_4 (7.10 g were dissolved in 250 mL of H_2O with 0.1 M NaCl) and a 0.1 M Citric acid solution (4.81 g were dissolved in 250 mL of H_2O) were prepared and filtered. Buffers of pH range from 3.0 to 8.0 were prepared by mixing this two stock solutions and adjusting the pH. Duplex DNA studies were performed at pH = 7.35 in the presence of 1.1 equivalents of the complementary strand. Before measurements, oligonucleotides were heated at 90 °C for 5 minutes, and slowly cooled to room temperature overnight.

7.4.5. Circular Dichroism (CD) Studies

CD melting experiments were carried out using a 0.1 cm path length thermo-controlled CD quartz cell. Spectra were collected at 25 °C between 220 and 340 nm with 0.1 nm steps, 2 nm band width, and a scanning rate of 50 nm/min. Three scans were measured and averaged for each reported spectrum. A concentration solution of 5 μM of pre-folded DNA was used for all experiments.

7.4.6. Thermal Denaturation Studies

CD melting experiments were carried out using a 0.1 cm path length thermo-controlled CD quartz cell. A concentration solution of 5 μM of pre-folded DNA were equilibrated for 15 minutes at 5 °C and slowly ramped to 95 °C at a rate of 10 °C/h. The molar ellipticity of each sample was monitored at $\lambda = 288$ nm for *i*-motif DNA and at $\lambda = 262$ nm for duplex DNA. The melting temperatures were determined by plotting the molar ellipticity of the sample versus the temperature using a dose-response non-linear regression. The average of the heating and cooling curves was taken for duplex DNA and only the heating curve was taken for *i*-motif DNA. The reproducibility of the measurements is within 0.3 °C.

7.5. Fluorescence in DNA

7.5.1. General

All fluorescence spectra were measured with pre-folded DNA in 96-well plate. Quantum yields of ^{DMA}**C** (**53**) in the context of oligonucleotides were determined as described above. The emission spectra of the probe in the context of the DNA were collected using the following wavelengths of excitation: λ_{ex} = 370 nm in single- and double-stranded DNA, λ_{ex} = 390 nm in *i*-motif DNA. The excitation spectra of the probe in the context of the DNA were collected using the following wavelengths of emission: λ_{em} = 535 nm in single- and double-stranded DNA and in *i*-motif DNA. ^{DMA}**C** (ϕ = 0.03 at pH = 7.35 with λ_{ex} = 370 nm and ϕ = 0.04 at pH = 5.50 with λ_{ex} = 410 nm) in H_2O was used as the fluorescent standard for the relative quantum yields (ϕ) of ^{DMA}**C** in DNA. All fluorescent spectra were measured using 4 µM of DNA in a buffer mixture of 20 mM of Na_2HPO_4 and 10 mM of citric acid supplemented with 10 mM NaCl at pH = 7.35 and 5.50. The relatively high absorbance of DNA samples at 260 nm (< 0.4 AU) caused a significant attenuation of the excitation beam that reduced the fluorescence intensities from the probe. This was evidenced by non-linear relationships between raw fluorescence intensities $F_{raw}(\lambda_{ex})$ and DNA concentrations above 50 nM when samples were photoexcited at 260 nm (Figure S8). $F_{raw}(\lambda_{ex})$ values at each excitation wavelength λ_{ex} were therefore corrected by multiplication with the following correction factor (CF) to obtain corrected $F(\lambda_{ex})$ values,[266] where $A(\lambda_{ex})$ is the raw absorbance of the sample at each wavelength of excitation:

$$CF = 2.303 * \frac{A_{(\lambda ex)}}{1 - 10^{-A(\lambda ex)}}$$

To check the validity of this approach, serial dilutions of ^{DMA}**C**-containing single-stranded, double-stranded and *i*-motif DNAs were analyzed for fluorescence intensity as a function of sample concentration (Figure S10.1). When the samples were excited at 260 nm, linear relationships were observed only after the application of the CF. At DNA concentrations of 50 nM and below, the fluorescence intensities for the corrected and uncorrected fluorescence measurements converged to the same values due to the low absorbance of the samples. These results validate the use of the CF. It should be emphasized that this CF is not needed for dilute DNA samples (< 50 nM) having minimal optical densities (OD < 0.01) at 260

nm. All reported excitation and emission spectra (λ_{ex} = 260 nm) as well as energy transfer efficiencies were corrected using CF.

Anisotropy measurements were performed with pre-folded DNA in 96-well plate and corrected for background and buffer contributions. Fluorescence anisotropy was calculated according to the following equation, where $I_{\parallel}$ and $I_{\perp}$ refers to the vertically and horizontally polarized light, respectively:

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$$

7.5.2. Fluorescence Intensity Correction Factor

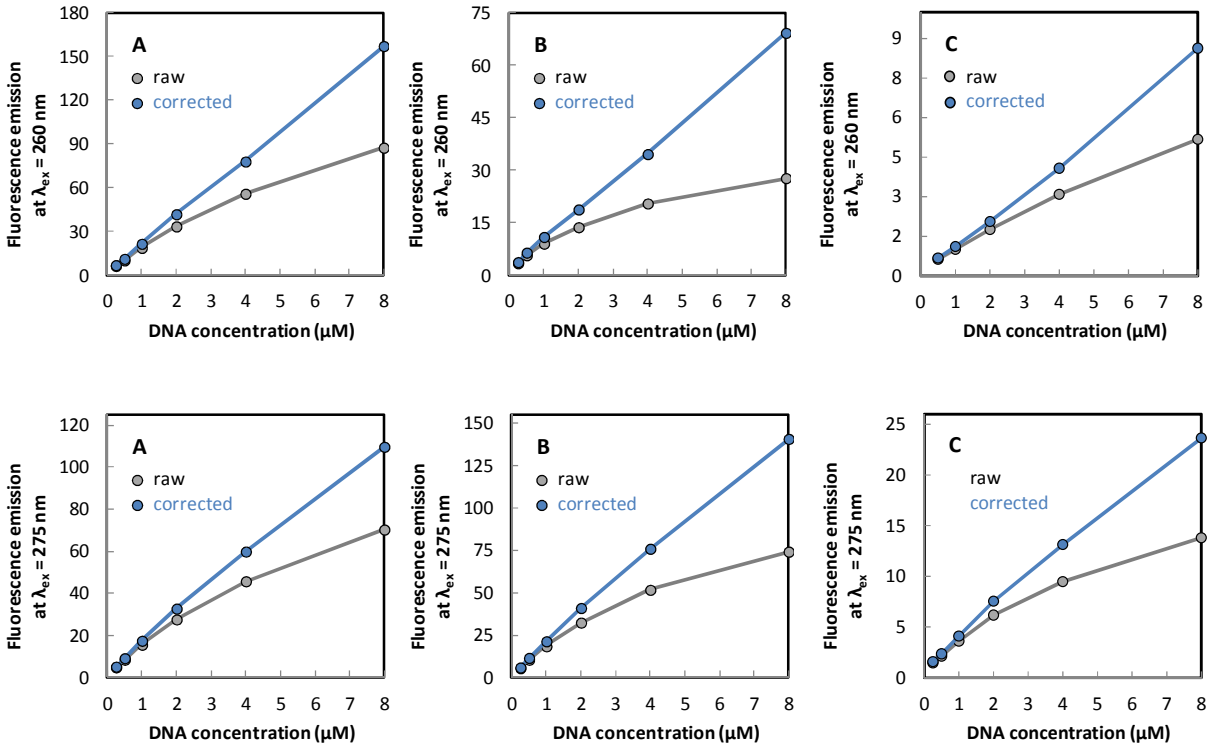


Figure 7.1. Sensitivity analysis and effect of the correction factor (CF). Raw (grey) and corrected (blue) fluorescence emission of ^{DMA}**C**-containing oligonucleotide prepared as (A) unfolded (B) duplex and (C) *i*-motif DNA at different concentration upon excitation at λ_{ex} = 260 nm (**TeloX2**) and at λ_{ex} = 275 nm (**TeloX15**) at pH = 7.35 and 5.50 with **X** = ^{DMA}**C**. Corrected values were obtained by the multiplication of the raw fluorescence areas or intensities by the correction factor (CF) accounting for the attenuation of the excitation beam and reduction of the fluorescence from absorption of the sample at each wavelength of excitation. All samples were prepared in a buffer mixture of 0.02 M of Na_2HPO_4 and 0.01 M of citric acid supplemented with 0.01 M NaCl.

7.5.3. Energy Transfer Efficiencies

Energy transfer efficiencies were calculated according to a modified reported method.[267] The transfer efficiency at any given excitation wavelength (λ_{ex}) is defined as ηt (λ_{ex}) as shown in the following equation:

$$\eta_t\,(\lambda_{ex}) = \frac{A_A\,(\lambda_{ex})}{A_D\,(\lambda_{ex})}\left(\frac{F\,(\lambda_{ex})}{F_A\,(\lambda_{ex})} - 1\right)$$

Where A_A (λ_{ex}) is the contribution of the acceptor (fluorescent probe ^{DMA}**C**) to the absorbance at wavelength λ_{ex}, A_D (λ_{ex}) is the contribution of the donor residues (unmodified DNA bases) to the absorbance at wavelength λ_{ex}. F (λ_{ex}) is the fluorescence intensity of the ^{DMA}**C**-modified oligonucleotide upon excitation at wavelength λ_{ex} and F_A (λ_{ex}) is the contribution of the acceptor (only) to the fluorescence intensity obtained upon excitation at wavelength λ_{ex}. Unless stated otherwise, λ_{ex} was set to λ_{ex} = 260 nm and fluorescence intensities were measured at $\lambda_{em} = \lambda_{max}$ = 515 – 550 nm in single- and double-stranded DNA and in *i*-motif DNA. The absorbance A_A (λ_{ex}) and fluorescence $F_{A,raw}$ (λ_{ex}) of the acceptor were obtained by measuring the absorption and emission spectra of nucleoside monomer solutions of the same concentration as the probe in DNA oligonucleotides. The actual contribution of the acceptor to the fluorescence intensity of ^{DMA}**C**-modified oligonucleotide, F_A (λ_{ex}), was calculated by multiplication of the measured fluorescence $F_{A,raw}$ (λ_{ex}) by the factor: ϕ (^{DMA}**C** in DNA) / ϕ (^{DMA}**C** monomer) to account for the differences in quantum yields of the probe in the context of DNA as compared to the monomer. Quantum yields (ϕ) were measured by direct excitation of the probe ^{DMA}**C** (ϕ = 0.03 at pH = 7.35 with λ_{ex} = 370 nm and ϕ = 0.04 at pH = 5.50 with λ_{ex} = 410 nm) as described above. The absorbance of the donor A_D (λ_{ex}) at wavelength λ_{ex} was calculated by subtracting the absorbance of the probe A_A (λ_{ex}) from the absorbance of the modified DNA oligonucleotides A (λ_{ex}). The measured fluorescence intensities F_{raw} (λ_{ex}) at each excitation wavelength λ_{ex} were corrected by the correction factor (CF) to get the corrected F (λ_{ex}) values. All measurements were performed with 4 μM solutions of pre-folded oligonucleotides prepared as described above.

7.6. Strand Displacement Assays

^{DMA}**C**-modified duplexes containing **C** and **G** oligonucleotides (4 μM) were prepared in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl) at pH = 7.35. The invading strand **I** (4.0 equiv., final concentration = 16 μM) was added at 25 °C, quickly mixed, and placed into the spectrometer to initiate the reaction. For the "simple" strand displacement assays conducted at pH = 7.35, the fluorescence anisotropy of the ^{DMA}**C**-modified oligonucleotide was measured at λ_{ex} = 370 nm and λ_{em} = 535 nm. For *i*-motif-mediated strand-displacement assays, the fluorescence intensity was measured at λ_{ex} = 370 nm and λ_{em} = 535 nm.

TH	(I) Match Sequence (5'-3')	(G) Match Sequence (3'-5')
1	CCC TAA CCC TAA CCC TAA CCC **C**	GGG ATT GGG ATT GGG ATT GGG **G**
2	CCC TAA CCC TAA CCC TAA CCC **TC**	GGG ATT GGG ATT GGG ATT GGG **AG**
3	CCC TAA CCC TAA CCC TAA CCC **ATC**	GGG ATT GGG ATT GGG ATT GGG **TAG**
4	CCC TAA CCC TAA CCC TAA CCC **TATC**	GGG ATT GGG ATT GGG ATT GGG **ATAG**
1	CCC TAA CCC TAA TCC TAA CCC **C**	GGG ATT GGG ATT AGG ATT GGG **G**
2	CCC TAA CCC TAA TCC TAA CCC **TC**	GGG ATT GGG ATT AGG ATT GGG **AG**

Table 7.5. Sequence composition of the invading (**I**) and complementary (**G**) oligonucleotides used for the DNA strand displacement reactions with ^{DMA}**C**- (top) and ^{DMA}**T**- (down) modified oligonucleotide.

TH	(I) Mutant Sequence (5'-3')	(G) Mutant Sequence (3'-5')
2	CTC TAA CCC TAA CTC TAA CCC **TC**	GAG ATT GGG ATT GAG ATT GGG **AG**
3	CTC TAA CCC TAA CCC TAA CTC **ATC**	GAG ATT GGG ATT GAG ATT GGG **TAG**

Table 7.6. Sequence composition of the invading mutant ($\mathbf{I}_{mut}$) and complementary (**G**) oligonucleotides used for the DNA strand displacement reactions with ^{DMA}**C**-modified oligonucleotide.

The strand displacement reaction is modeled as a simple bimolecular reaction (eq. 1); where **GC** represents the initial duplex, **I** the incoming strand, **GI** the newly formed duplex, **C** the displaced ^{DMA}**C**- or ^{DMA}**T**-containing strand, k_1 and $k_{1'}$ are the forward and backward rate constants:[268-270]

$$\mathbf{GC} + \mathbf{I} \underset{k_{1'}}{\overset{k_1}{\rightleftharpoons}} \mathbf{GI} + \mathbf{C}$$

Given the higher kinetic and thermodynamic stability of **GI** versus **GC**, the back reaction $k_{1'}$ is neglected. The strand displacement reaction can therefore be given as a function of time (t) (eq. 2):

$$\frac{[\mathbf{C}]}{[\mathbf{GC}]_0} = \frac{k_1[\mathbf{I}]_0 t}{1 + k_1[\mathbf{I}]_0 t}$$

The concentration of the incoming strand ($[\mathbf{I}]_0$ = 16 μM) is in large excess as compared to the initial duplex ($[\mathbf{GC}]_0$ = 4 μM). The time-dependent displacement reaction is therefore approximated by (eq. 3):

$$\frac{[\mathbf{C}]}{[\mathbf{GC}]_0} = 1 - e^{-k_1[\mathbf{I}]_0 t}$$

Accordingly, the "simple" strand displacement reactions monitored using fluorescent anisotropy data were fit to the monoexponential equation (eq. 4); where the rate constant k_1 is obtained by fitting the anisotropy (r) versus time (t):

$$r = e^{-k_1[\mathbf{I}]_0 t}$$

Strand displacement reactions involving i-motifs were monitored using fluorescent intensity (f) and fit to a biexponential equation (eq. 5); where the rate constant k_1 and k_2 are obtained by fitting the fluorescence intensity (f) versus time (t), and Fr is the fractional contribution:

$$f = Fr_1 \cdot e^{-k_1[\mathbf{I}]_0 t} + Fr_2 \cdot e^{-k_2[\mathbf{I}]_0 t}$$

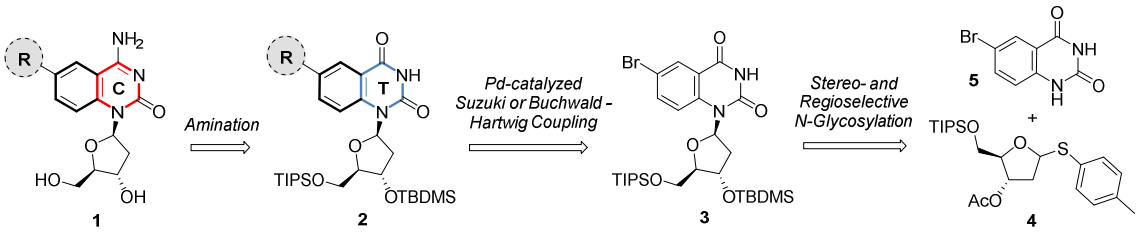
Chapter 8 | Summary & Conclusions

8.1. Summary

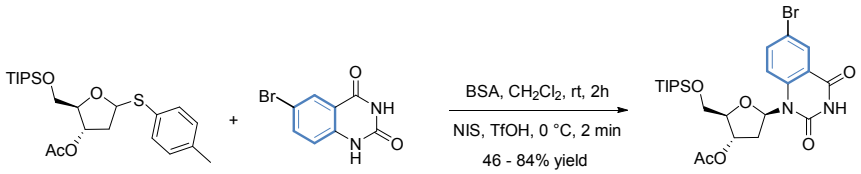
Fluorescent Nucleosides for Probing Proton-Coupled DNA Folding in Real Time

The transfer of genetic information from nucleotide sequences to proteins is the central defining feature of life on Earth. With the vast majority of the human genome composed of non-coding sequences, DNA has been regarded by most as a uniform double helix and a passive library of genetic information. However, DNA is a highly dynamic molecule and non-duplex DNA structures play important roles in the regulation of gene expression and chromosome stability. Cytosine-rich sequences can self-assemble into four-stranded structures called *i*-motifs that contain hemi-protonated C-C$^+$ base pairs. The presence of this unique architecture *in vivo* is still a source of debate, since *i*-motif folding requires slightly acidic conditions (pH ≤ 6) to promote protonation of the cytosine residues.

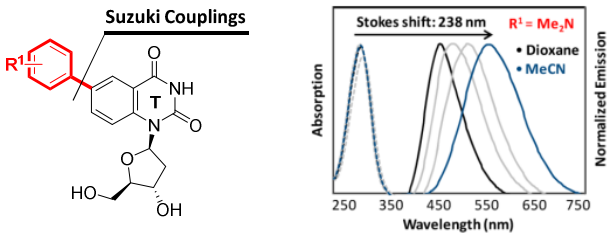
To detect *i*-motif structures *in vivo*, we envisioned the utilization of a minimally disruptive fluorescent nucleoside analog. For this purpose, we designed and evaluated a new family of fluorescent cytosine (**1**) and thymidine (**2**) nucleosides containing a quinazoline fluorophore. To create "push-pull" fluorophores, electron-donating groups were added to position *C6* of quinazoline, which was selected for its negligible impact on the physicochemical properties of *N3* due to the lack of conjugation. The targeted compounds were prepared using Pd-catalyzed Suzuki-Miyaura or Buchwald-Hartwig cross-couplings reactions, and stereoselective *N*-glycosylation for the assembly of the nucleobase (**5**) and the furanose ring (**4**).



To facilitate synthesis of the target compounds, an efficient method for the *N*-2-deoxyribosylation of modified nucleobases by thioriboside donors was developed. In this approach, thioethers are selectively activated by iodination to produce five-membered ring oxocarbenium ions. The conformation preferences of these intermediates were tuned using various functional groups on the furanose ring to induce stereochemistry at C1 (α/β = 1.0:4.0 to 4.5:1.0) upon condensation with an activated nucleobase. This strategy enabled the stereo- and regioselective synthesis of various pyrimidine nucleosides, including ring-expanded pyrimidine derivatives containing sulfur and bromine atoms.

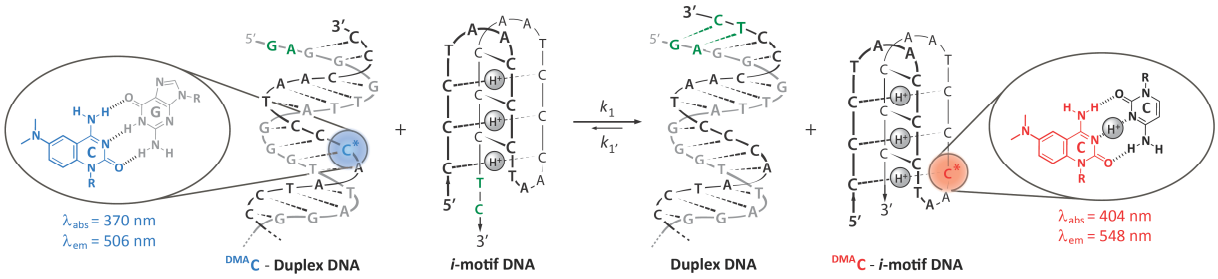


Following *N*-glycosylation, Suzuki cross-coupling reactions were used to prepare a series of biaryl fluorophores that exhibit charge transfer from the benzene ring to the pyrimidine group. The magnitude of this "push-pull" system was augmented by the presence of electron-donating substituents (R^1). The donor-acceptor properties of these fluorophores was demonstrated by a linear correlation of their HOMO-LUMO energy gaps to emission wavelengths. The exceptional solvatochromic properties of a dimethylaniline derivative revealed the presence of a twisted intramolecular charge transfer (TICT) emissive state.



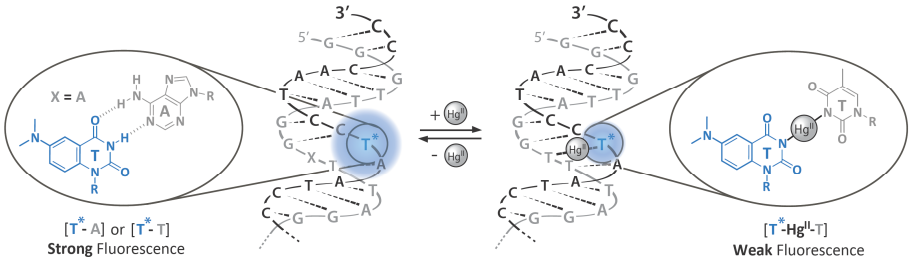
Among the synthesized fluorophores, derivatives having a dimethylaniline unit fused to the pyrimidine core, ^{DMA}**C** and ^{DMA}**T**, were recognized as promising candidates for incorporation into DNA. The corresponding phosphoramidite of ^{DMA}**C** was synthesized and incorporated into the cytosine-rich

human telomeric repeat sequence that folds into *i*-motif structures under slightly acidic conditions (pH = 5.75). Due to its effective mimicry of cytidine, ^{DMA}**C** has little or no impact on DNA folding or stability. Notably, ^{DMA}**C** exhibited higher fluorescence quantum yields (ϕ) in the context of nucleic acids (ϕ = 0.02 – 0.11), as compared to the free nucleoside in water (ϕ = 0.03). Folding into *i*-motif structures with concomitant protonation at *N3* resulted in red-shifted excitation and emission maxima. ^{DMA}**C**'s fluorescence properties were also highly sensitive to local conformational dynamics such as fluorescence anisotropy (r = 0.04 – 0.27), and global properties such as nucleobase-nucleobase energy transfer efficiency (ηt = 0.03 – 0.17). The parameters were utilized for tracking conformational changes in real time using DNA strand displacement assays. These were conducted by mixing ^{DMA}**C**-labeled duplexes containing a 5' single-stranded overhang with an excess of unlabeled DNA to initiate thermodynamically favorable unfolding-refolding that releases the ^{DMA}**C**-labeled strand from its complement. Strand displacement reactions using folded *i*-motif DNA revealed that despite their low thermal stabilities, *i*-motifs can pose exceptionally high kinetic barriers to duplex formation (k = 1.0 $M^{-1}s^{-1}$, $t_{½} \approx$ 12 hours), under near-physiological conditions (pH = 5.75). These results suggest that energy-dependent processes might be needed to resolve *i*-motif structures formed *in vivo*.



Mercury is a highly cytotoxic and mutagenic heavy-metal that exhibits strong interactions with nucleic acids *via* specific binding to thymidine-thymidine (T–T) mismatches in duplex DNA. To probe the coordination of Hg^{II} ions to thymidine in T–Hg^{II}–T base pairs, a fluorescent thymidine analog ^{DMA}**T** was introduced into DNA using phosphoramidite chemistry. ^{DMA}**T** exhibited very similar conformational and physicochemical properties as native thymidine. Upon addition of Hg^{II} ions to DNA containing ^{DMA}**T**–T base pairs, ^{DMA}**T**-modified duplexes exhibited increased thermal melting temperatures comparable to A–T base pairs. The resulting site-specific formation of ^{DMA}**T**–Hg^{II}–T was monitored using the photophysical properties of ^{DMA}**T**. Among many metals ions evaluated, ^{DMA}**T**'s fluorescence showed

high selectivity for Hg^{II} ions and could therefore be used as a specific site-selective sensor for this highly toxic heavy-metal.



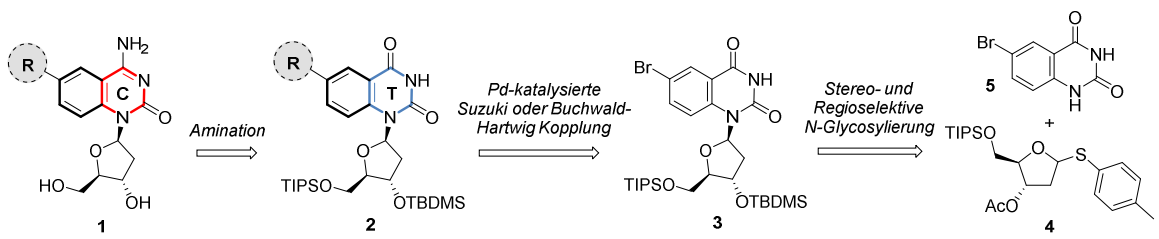
The combination of cytosine and thymidine nucleobases with a dimethylaniline unit has provided highly sensitive fluorescent probes. Well resolved and characteristic fluorescent properties, such as emission spectra, fluorescence anisotropy and energy transfer efficiencies afford a sensitive method for the direct readout of folded structures. ^{DMA}**C** was the first nucleoside analog designed for detecting proton-coupled DNA folding and it revealed unexpected slow unfolding of *i*-motif DNA. ^{DMA}**T** provided the first thymidine mimic capable of reporting site-selective incorporation of Hg^{II} ions in duplex DNA. With such powerful properties, we believe that the utilization of ^{DMA}**C** and ^{DMA}**T** in different systems will facilitate the exploration of other biological processes and thereby contribute to a better understanding of nature.

8.2. Zusammenfassung

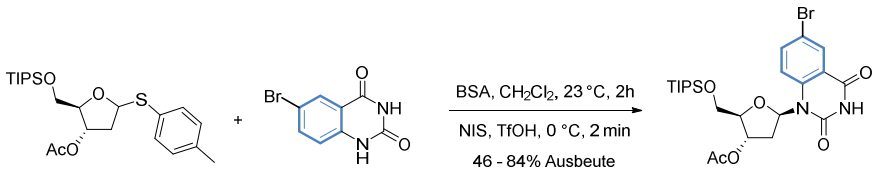
Fluoreszierende Nukleoside um Proton-Gekoppelte DNA-Faltung in Echtzeit zu Verfolgen

Die Übersetzung der genetischen Information von der Nukleotidsequenz der DNA in Proteine ist ein zentrales Kennzeichen des Lebens auf der Erde. Da der überwiegende Anteil des menschlichen Genomes aus nicht-kodierenden Sequenzen besteht, wurde DNA lange Zeit als gleichmässiger Doppelhelix sowie als passive Sammlung genetischer Informationen gehalten. DNA ist jedoch hoch dynamisch wobei Strukturen abgesehen vom Doppelhelix eine wichtige Rolle in der Regulierung von Genexpression und Chromosomenstabilität spielen. Die cytosinreichen DNA Sequenzen können sich eigenständig in vier-strängige Strukturen aus halb-protonierten C-C$^+$ Basenpaaren, sogenannten *i*-Motiv DNA, anordnen. Ob die einzigartige Architektur des i-Motivs *in vivo* vorkommt, steht dauernd zur Debatte, da es zur Faltung zum *i*-Motiv leicht saure Bedingungen (pH < 6) benötigt um die Protonierung des Cytosine zu ermöglichen.

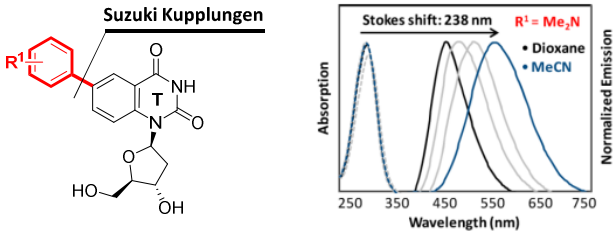
Um *i*-Motiv Strukturen *in vivo* nachzuweisen, haben wir ein fluoreszierendes Nukleosidanalog ausgedacht, welcher die sekundäre Struktur kaum verändert. Zu diesem Zweck wurde eine neue fluoreszierende Familie von Cytosin- (**1**) und Thymidin- (**2**) Nukleosiden mit Quinozalin-Fluorophoren entwickelt und untersucht. Um ein „push-pull" Fluorophoresystem zu schaffen, wurden elektronengebende Gruppen auf Position *C6* der Quinazoline eingefügt. Die C6 Position wurde gezielt ausgewählt, da aufgrund mangelnder Konjugation die Auswirkungen auf die physikalisch-chemische Eigenschaften von *N3*-Position minimal sind. Die gewünschten Verbindungen wurden mittels stereoselektiven Glykosylierung, um die Nukleinbase (**5**) mit dem Furanosering (**4**) zu verbinden, gefolgt von Pd-katalysierter Suzuki-Miyaura oder Buchwald-Hartwig Kupplungsreaktionen, synthetisiert.



Um die Synthese der gewünschten Hauptverbindungen zu vereinfachen, wurde eine effiziente Methode zur *N*-2-Deoxyribosylierung der modifizierten Nukleobasen mit einem Thioribosiddonator entwickelt. Thioether können selektiv aktiviert werden, da sie in Gegenwart eines Iodoniums ein fünfringes Oxocarbenium Ion bilden. Die bevorzugte Konformationen der Zwischenprodukte wurde mittels verschiedener funktionaler Gruppen am Furanosering abgestimmt. Dadurch konnte die Stereochemie auf *C1* nach zugefügter, aktivierter Nukleobase in einen Verhältnis von (α/β = 1.0:4.0 to 4.5:1.0) nachgewiesen werden. Diese Strategie ermöglicht eine stereo- und regioselektive Synthese von verschiedenen Pyrimidine-Nukleosiden, einschliesslich ausgedehnten Pyrimidinring-Derivaten, die Schwefel- und Bromatome enthalten.

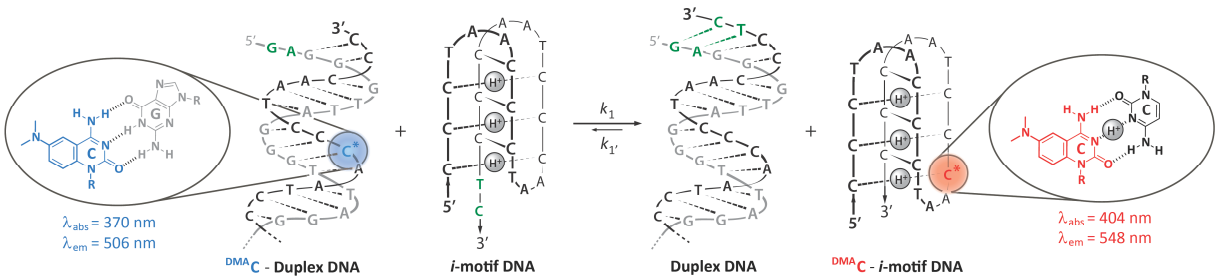


Nach erfolgter *N*-Glykoliserung wurden durch Suzuki-Miyaura Kupplungsreaktionen verschiedene „biaryl" Fluorophore synthetisiert, die Ladungsverschiebungen vom Benzolring zur Pyrimidingruppe aufweisen. Die Stärke dieses „push-pull" Systemes wurde durch die Präsenz von elektronengebenden Substituenten (R^1) erhöht, was durch die lineare Korrelation der Energiedifferenz zwischen HOMO und LUMO, und der Emissionswellenlängen nachgewiesen werden konnte. Grund für die aussergewöhnlichen solvatochromen Eigenschaften von Dimethylaniline-substituierten Derivate ist das Vorhandensein des „twisted intramolecular charge transfer" (TICT) Emmisionszustand.



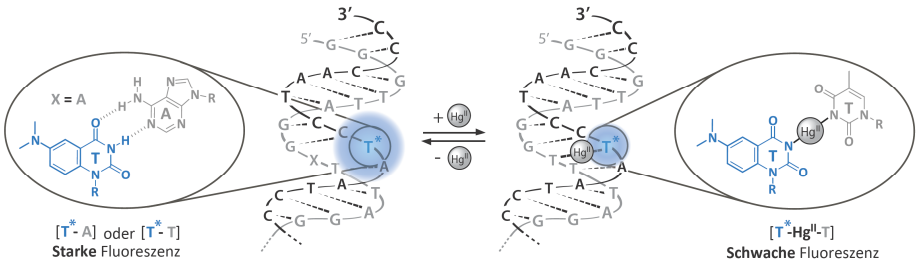
Unter den verschiedenen synthetisierten Fluorophoren wurden die Dimethylaniline-substituierten Pyrimidin-Einheiten ^{DMA}**C** und ^{DMA}**T** als vielversprechende Kandidaten für den Einbau in die DNA erkannt. Dazu wurde das zugehörige Phosphoramidit ^{DMA}**C** synthetisiert und in die cytosinreiche

menschliche telomerische Wiederholungssequenz eingebaut, da sich diese Sequenz unter leicht sauren Bedingungen (pH =5.75) zum *i*-Motiv Struktur faltet. Dank der effektiven Mimik von Cytidin, hat ^{DMA}**C** wenig oder keine Auswirkungen auf die Faltung oder Stabilität der DNA. Zudem weist ^{DMA}**C** höhere Fluoreszenzquantenausbeuten (ϕ) in der Umgebung von Nukleinsäuren (ϕ = 0.02 – 0.11), als im freien Nukleosid in Wasser (ϕ = 0.03) auf. Die Faltung zur *i*-Motiv Struktur mit gleichzeitige Protonierung von *N3* wurde durch rot-verschobende Anregungs- und Emissionsmaxima nachgewiesen. Die Fluoreszenzeigenschaften von ^{DMA}**C** sind ebenfalls hochsensibel gegenüber lokaler Komformation-dynamik, wie zum Beispiel Fluoreszenz-Anisotropie (r = 0.04 – 0.27), sowie globale Eigenschaften wie die Energietransfer-Effizienz (ηt = 0.03 – 0.17). Dank dieser Parameter konnte die Konformationsänderung bei DNA-Strang-Verdrängungsassays in Echtzeit verfolgt werden. Dabei wurden ^{DMA}**C**-markierte Duplexe mit einem überständigen 5'-Einzelstrang mit einem Überschuss an unmarkierter DNA gemischt. Dies löste einen thermodynamisch begünstigten Entfaltungs-Faltungsprozess aus, wobei der ^{DMA}**C**-markierte Strang von seinem komplementären Strang befreit wird. Strang-Verdrängungsassays mit gefallteter *i*-Motiv DNA zeigen, dass *i*-Motive unter fast physiologischen Bedingungen (pH = 5.75) trotz niedriger thermalen Stabilität ausserordentlich hohe kinetische Barrieren zur Duplexbildung aufweisen (k = 1.0 $M^{-1}s^{-1}$, $t_{½}$ ≈ 12 Stunden). Diese Resultate zeigen, dass energieabhängige Prozesse nötig sind, um *i*-Motive *in vivo* nachzuweisen.



Quecksilber ist ein hoch zytotoxisches und mutagenes Schwermetall, welches durch spezifische Bindung zu Thymidin-Thymidin (T–T) Fehlpaaren in der Duplex-DNA starke Interaktionen mit Nukleinsäuren aufweist. Um die Koordination von Hg^{II} Ionen zu Thymidin im T–Hg^{II}–T Basenpaar zu erkennen, wurde ein fluoreszierendes Thymidinanalog ^{DMA}**T** in DNA mittels Phophoramidit-Chemie eingefügt. ^{DMA}**T** hat eine sehr ähnliche Konformation sowie physikalisch-chemische Eigenschaften wie natürliches Thymidin. Bei Zugabe von Hg^{II} Ionen zu DNA mit ^{DMA}**T**–T Basenpaaren zeigte sich, dass die

geänderte ^{DMA}**T** Duplex erhöhte Schmelztemperaturen im Vergleich zu A–T Basenpaaren aufweist. Diese spezifische Bildung von ^{DMA}**T**–Hg^{II}–T wurde durch photophysikalische Eigenschaften von ^{DMA}**T** beobachtet. Unter vielen getesteten Metallionen zeigt die Fluoreszenz des ^{DMA}**T** hohe Selektivität für Hg^{II} Ionen. Somit kann ^{DMA}**T** auch als spezifischer und selektiver Sensor für dieses hoch toxische Schwermetall benutzt werden.



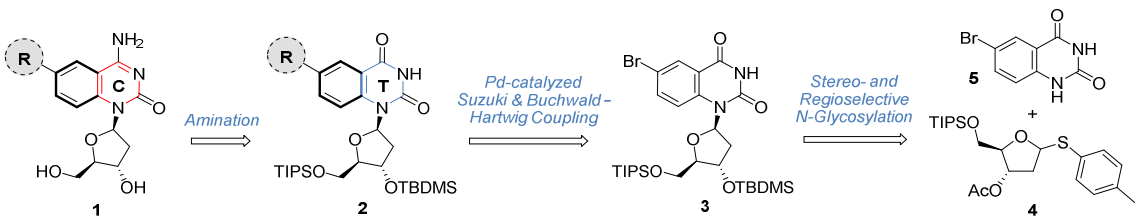
Die Kombination von Cytosin- und Thymidin-Nukleobasen, welche mit Dimethylaniline substituiert sind, hat zu hochsensitive Fluoreszenzmarker geführt. Gut aufgelöste und charakteristische Fluoreszenzeigenschaften wie Emissionspektren, Fluoreszenz-Anisotropie und Energietransfer-Effizienz ergeben eine empfliche Methode zur direkten Detektion gefalteter Strukturen. ^{DMA}**C** ist das erste designte Nukleosidanalogen zur Erkennung von Proton-gekoppelter DNA-Faltung. Zudem konnte wir mittels ^{DMA}**C** nachweisen, dass die Entfaltung des *i*-Motiv unerwartet langsam abläuft. ^{DMA}**T** ist die erste fluoreszierende Thymidin-Mimik, womit der positionsgenaue Einbau von Hg^{II} Ionen in Duplex-DNA festgestellt werden kann. Mit diesen überragenden Eigenschaften glauben wir, dass die Nutzung von ^{DMA}**C** und ^{DMA}**T** in verschiedenen Systemen die Erforschung anderer biologischer Prozesse vereinfacht und dadurch zu einem besseren Verständnis der Natur führt.

8.3. Résumé

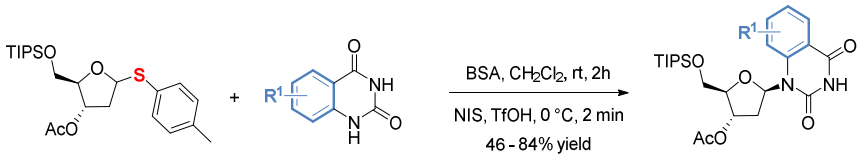
Nucléosides Fluorescents Pour Détecter le Dynamisme de l'ADN en Temps Réel

Le transfert de l'information génétique à partir d'une séquence de nucléotides en protéine définit l'importance de la vie sur terre. Avec la majorité du génome humain composée essentiellement de séquences non-codantes, l'ADN a été longtemps perçu comme une double hélice uniforme et une bibliothèque passive d'information génétique. L'ADN est, cependant, une molécule très dynamique et les structures en dehors de la double hélice jouent d'importants rôles dans la régulation de l'expression des gènes et dans la stabilité des chromosomes. Les séquences riches en cytosine peuvent se replier en structure de quatre brins composée de dimer de cytosines protonées C-C$^+$, que l'on appelle *i*-motif. La présence de cette architecture unique *in vivo* fait toujours l'objet de débats, puisque le repliement de l'*i*-motif nécessite des conditions légèrement acides (pH ≤ 6) pour générer la protonation des résidus de cytosine.

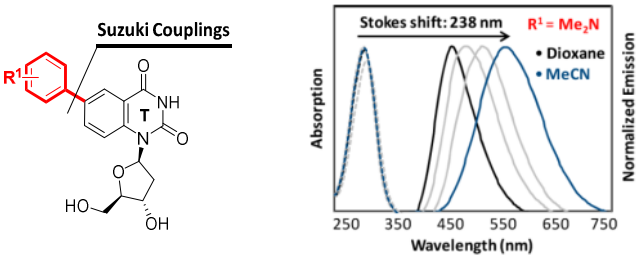
Afin de détecter les structures d'*i*-motif *in vivo*, nous avons envisagé d'utiliser un nucléoside analogue fluorescent minimisant la dénaturalisation de la structure. Dans ce but, nous avons conçu et évalué une nouvelle famille fluorescente de cytosines (**1**) et de thymidines (**2**) composée d'un noyau quinazoline. Afin de créer un système fluorescent de type "push-pull", nous avons ajouté des groupements donneurs d'électrons en position *C6* de la quinazoline. Cette position a été sélectionnée pour négliger l'effet sur les propriétés physico-chimiques de la position *N3* par manque de conjugaison directe. Les composés cibles ont été préparés par couplages pallado-catalysés de Suzuki-Miyaura et de Buchwald-Hartwig, et par *N*-glycosylation stereoselective pour l'assemblage de la nucleobase avec le cycle furanose.



Afin de faciliter la synthèse des composés cibles, une méthode efficace de *N*-2-deoxyribosylation pour le couplage de nucleobases modifiées avec des thioribosides donneurs a été développée. Les thioethers peuvent être activés sélectivement en présence d'iodonium pour former des ions oxocarbeniums de cycle à cinq chaînons. Les conformations préférentielles de ces intermédiaires ont été contrôlées en variant les groupements fonctionnels du furanose afin de générer la stéréochimie en *C1* (α/β = 1.0:4.0 to 4.5:1.0) lors de la condensation d'une nucleobase activée. Cette stratégie a permis de réaliser la synthèse stéréo- et regioselective de divers nucléosides à base pyrimidine, comprenant des dérivés pyrimidines à cycle étendu composés d'atomes de soufre et de brome.

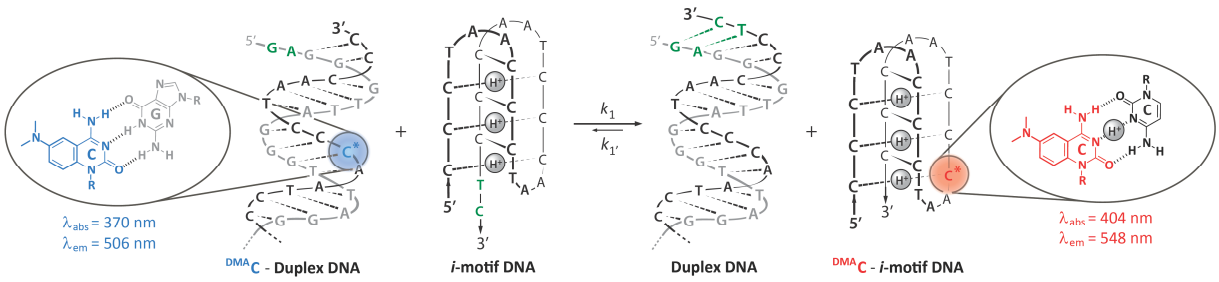


Après la *N*-glycosylation, des réactions de couplages croisés de Suzuki ont été réalisées pour préparer des fluorophores biaryls. Ces derniers entraînent un transfert de charge du cycle benzénique au groupement pyrimidine. L'introduction de groupement donneur d'électrons (R^1) a permis d'augmenter l'ampleur de ce système de "push-pull". Les propriétés donneur-accepteur de ces fluorophores ont été démontrées par la relation entre leurs niveaux d'énergie des orbitales frontières HOMO-LUMO et leurs longueurs d'ondes d'émission. Les propriétés solvatochromiques du dérivé possédant un groupement dimethylaniline ont révélé la présence de "twisted intramolecular charge transfer" (TICT) des états émissifs.



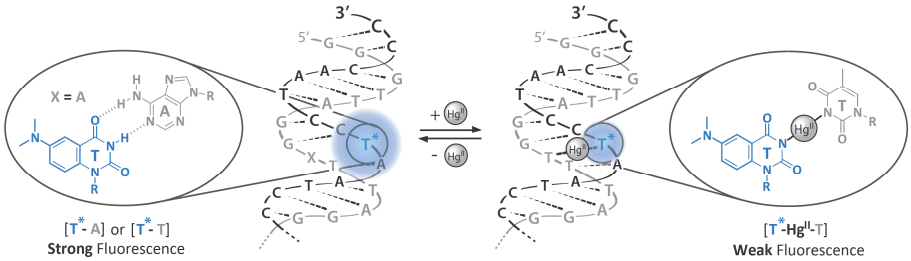
Parmi la famille de fluorophores synthétisée, les dérivés ayant un groupement dimethylaniline fusionné à la base pyrimidine, ^{DMA}**C** and ^{DMA}**T**, ont été sélectionnés. Le dérivé phosphoramidite correspondant à ^{DMA}**C** a été synthétisé et incorporé dans la séquence d'ADN du télomère humain qui se replie en *i*-motif en condition légèrement acide (pH = 5.75). ^{DMA}**C** imite la cytidine naturelle, et n'a donc

que très peu ou pas du tout d'impact sur le repliement et la stabilité de l'ADN. Notamment, ^{DMA}**C** possède un rendement quantique de fluorescence (ϕ) supérieur dans les acides nucléiques (ϕ = 0.02 – 0.11), par rapport au nucléoside libre (ϕ = 0.03). Le repliement en *i*-motif, combiné avec la protonation de la position *N*3, conduit au déplacement des longueurs d'ondes d'excitation et d'émission vers le rouge. Les propriétés fluorescentes de ^{DMA}**C** sont très sensibles au dynamisme local, telle que l'anisotropie de fluorescence (r = 0.04 – 0.27), ainsi qu'aux propriétés globales telle que l'efficacité de transfert d'énergie entre les nucleobases (ηt = 0.03 – 0.17). Ces paramètres ont ensuite été utilisés pour suivre en temps réel des réactions de déplacements de brins d'ADN. Celles-ci ont été réalisées en mélangeant un double brin marqué de ^{DMA}**C** avec un excès d'ADN non-marqué pour initier une réaction favorable de dépliage qui libère le brin marqué de ^{DMA}**C** de son complément. Les réactions de déplacements de brins d'ADN utilisant des structures repliées en *i*-motif de stabilités thermiques faibles, ont révélé qu'ils peuvent poser des barrières cinétiques exceptionnellement élevées à la formation de double brin (k = 1.0 $M^{-1}s^{-1}$, $t_{½} \approx$ 12 heures), en conditions presque physiologique. Ces résultats suggèrent que la combinaison de sources d'énergie extérieure, telle que l'activité des polymérases, sont nécessaires pour déterminer les structures d'*i*-motif *in vivo*.



Le mercure est un métal lourd très toxique et mutagène qui possède de fortes interactions avec les acides nucléiques en formant des liaisons spécifiques avec les mésappariements thymidine-thymidine (T–T) dans les doubles brins d'ADN. Afin de détecter la coordination des ions Hg^{II} avec la thymidine, une thymidine analogue fluorescente ^{DMA}**T** a été introduite dans l'ADN en utilisant la synthèse phosphoramidite. ^{DMA}**T** possède des propriétés physico-chimiques et conformationelles très similaires à la thymidine naturelle. Lors de l'ajout d'ions Hg^{II} en présence de mésappariements ^{DMA}**T**–T, les températures de dénaturation des oligonucléotides contenant ^{DMA}**T** atteignent des températures similaires aux appariements A–T. Cette formation sélective d'appariement ^{DMA}**T**–Hg^{II}–T a été suivie en utilisant les

différences de rendement de fluorescence par rapport aux appariements ^{DMA}**T**–T et ^{DMA}**T**–A. La fluorescence de ^{DMA}**T** a montré une grande sélectivité pour les ions Hg^{II}, par rapport à d'autres ions métalliques évalués, et peut donc être utilisé comme détecteur spécifique de ce métal très toxique.



La modification des bases cytosine et thymidine avec un simple groupement dimethylaniline a généré des molécules fluorescentes très sensibles. Les propriétés fluorescentes caractéristiques, tels que les spectres d'émission, l'anisotropie de fluorescence et l'efficacité d'énergie de transfert, ont permis le développement d'une méthode sensible pour détecter le repliement des structures d'ADN. ^{DMA}**C** a été le premier nucléoside analogue conçu pour détecter le repliement d'ADN couplés aux protons et a révélé des propriétés fondamentales inattendues de l'ADN *i*-motif. ^{DMA}**T** a fourni la première mimique de la thymidine pour détecter l'incorporation sélective des ions Hg^{II} dans l'ADN double brins. Avec de telles propriétés, nous sommes confiants que l'utilisation de ^{DMA}**C** et ^{DMA}**T** dans d'autres systèmes facilitera l'exploration d'autres processus biologiques et contribuera ainsi à une meilleure compréhension de notre nature.

List of Abbreviations and Symbols

A	Adenine
Ac	Acetyl
API	Atmospheric pressure ion
Aq.	Aqueous
BCL2	B-cell lymphoma 2
BDF	Base-discriminating fluorescent
Bz	Benzoyl
BSA	*N,O*-bis(trimethylsilyl)acetamide
C	Cytosine
MYC	Myelocytomatosis
CD	Circular dichroism
CF	Correction factor
COSY	Correlation spectroscopy
DAPI	4',6-Diamidino-2-phenylindole
Dba	Dibenzylideneacetone
DCM	Dichloromethane
DCE	1,2-Dichloroethane
DCI	*N,N*'-Diisopropylcarbodiimide
DFT	Density functional theory
DIPEA	*N,N*-Diisopropylethylamine
DMA	Dimethylaniline
DMAP	*N,N*-Dimethyl-4-aminopyridine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DMT	Dimethoxytrityl
DNA	2'-Deoxyribonucleic acid
Dppf	1,1'-Bis(diphenylphosphino)-ferrocene
ESI	Electrospray ionization
FRET	Förster resonance energy transfer
G	Guanine
Gly	Glycine
HPLC	High performance liquid chromatography
HR	High resolution
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HMDS	Hexamethyldisilazane
HOMO	Highest occupied molecular orbital
***i*Pr**	Isopropyl
IR	Infrared spectroscopy
LUMO	Lowest unoccupied molecular orbital
MS	Mass spectrometry
MS 4Å	Molecular sieves 4Å
NBS	*N*-Bromosuccinimide
NIS	*N*-Iodosuccinimide
NMR	Nuclear magnetic resonance
NISac	*N*-iodosaccharin
2PyG	8-(2-Pyridyl)-2'-deoxyguanosine
PDB	Protein data bank
PMB	*p*-Methoxybenzoyl
Pin	Pinacol
Ppm	Parts per million
QTOF	Quadrupole time-of-flight
RET	Rearranged during transfection
Rb	Retinoblastoma
RNA	Ribonucleic acid
ROESY	Rotating frame nuclear Overhauser effect spectroscopy
RT	Room temperature
Sat.	Saturated
SN	Nucleophilic substitution
SNP	Single nucleotide polymorphism
T	Thymidine

TBDMS	*tert*-Butyldimethylsilyl
TBAF	Tetra-*n*-butylammonium fluoride
TDS	Thermal difference spectrum
TEAA	Triethylammonium acetate
Telo	Hunan telomeric DNA sequence
TFA	Trifluoroacetic acid
TfOH	Triflic acid
TH	Toehold domain
THF	Tetrahydrofuran
TICT	Twisted intramolecular charge-transfer
TIPS	Triisopropylsilyl
TIPDS	1,1,3,3-tetraisopropyldisiloxyl
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
TLC	Thin layer chromatography
TPS	2,4,6-Triisopropylbenzenesulfonyl
Tol	Toluoyl
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
Wt	Wild-type

r	Fluorescence anisotropy
ϕ	Quantum yield
n	Refractive index
ηt	Energy transfer
λ_{ex}	Excitation wavelength
λ_{em}	Emission wavelength
λ_{abs}	Absorbance wavelength
E_T^{30}	Solvent polarity parameter
N	Nucleophilicity parameter
k	Rate constant
$t_{½}$	Half-life
ε	Molar extinction coefficient
δ	Chemical shift
J	Coupling constant
T_m	Thermal melting temperature

Appendixes

A1. Fluorescence Spectra of Six-Membered-Ring Biaryl Series

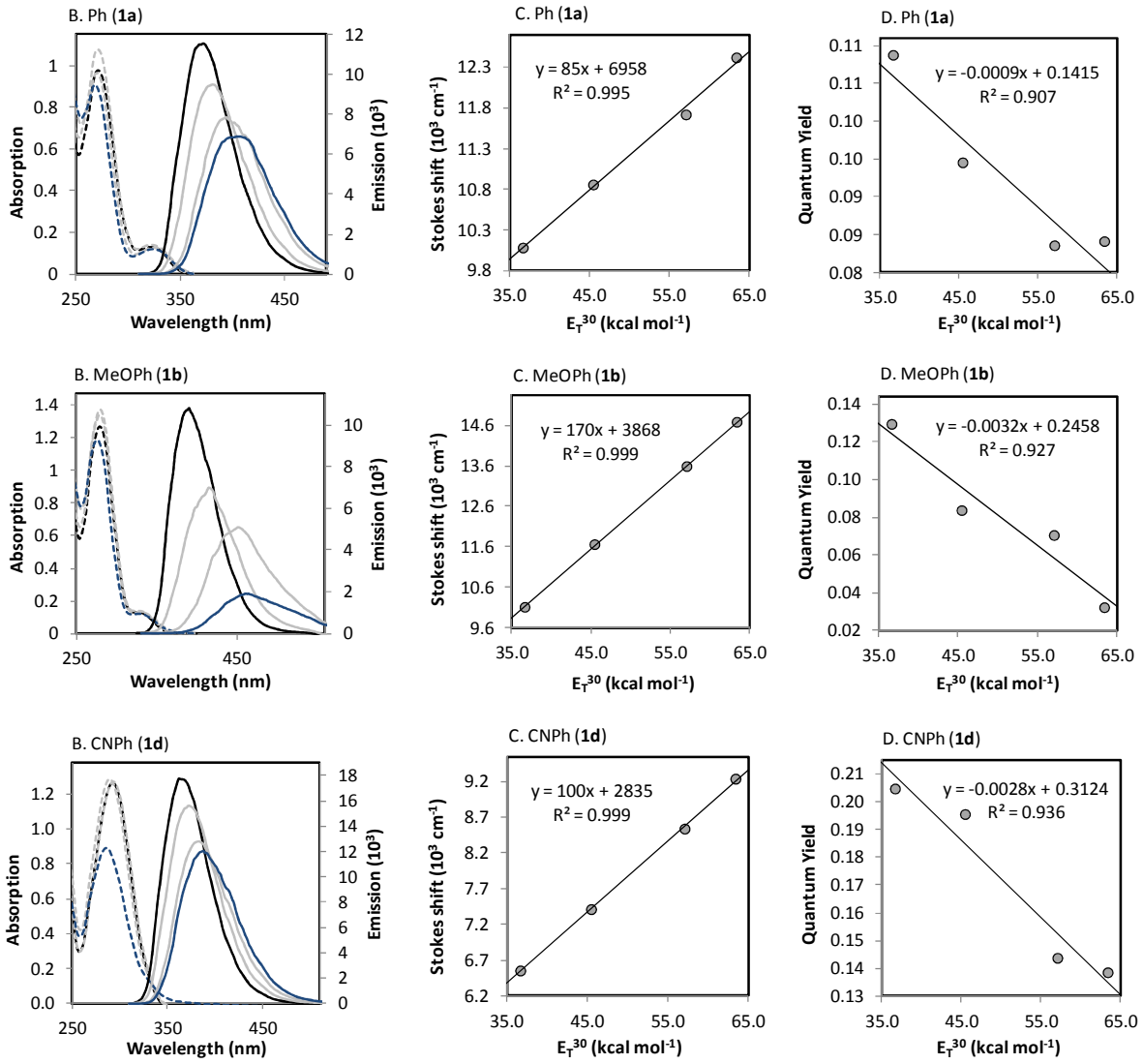


Figure A1. (B) Absorption (dashed) and emission (solid) spectra in water (blue) and dioxane (black) and mixtures (grey) at 40 μM. (C) Linear relationship between stokes shift and solvent polarity in dioxane, water and mixtures. (D) Linear relationship between quantum yield and solvent polarity in dioxane, water and mixtures.

A2. Fluorescence Spectra of Five-Membered-Ring Biaryl Series

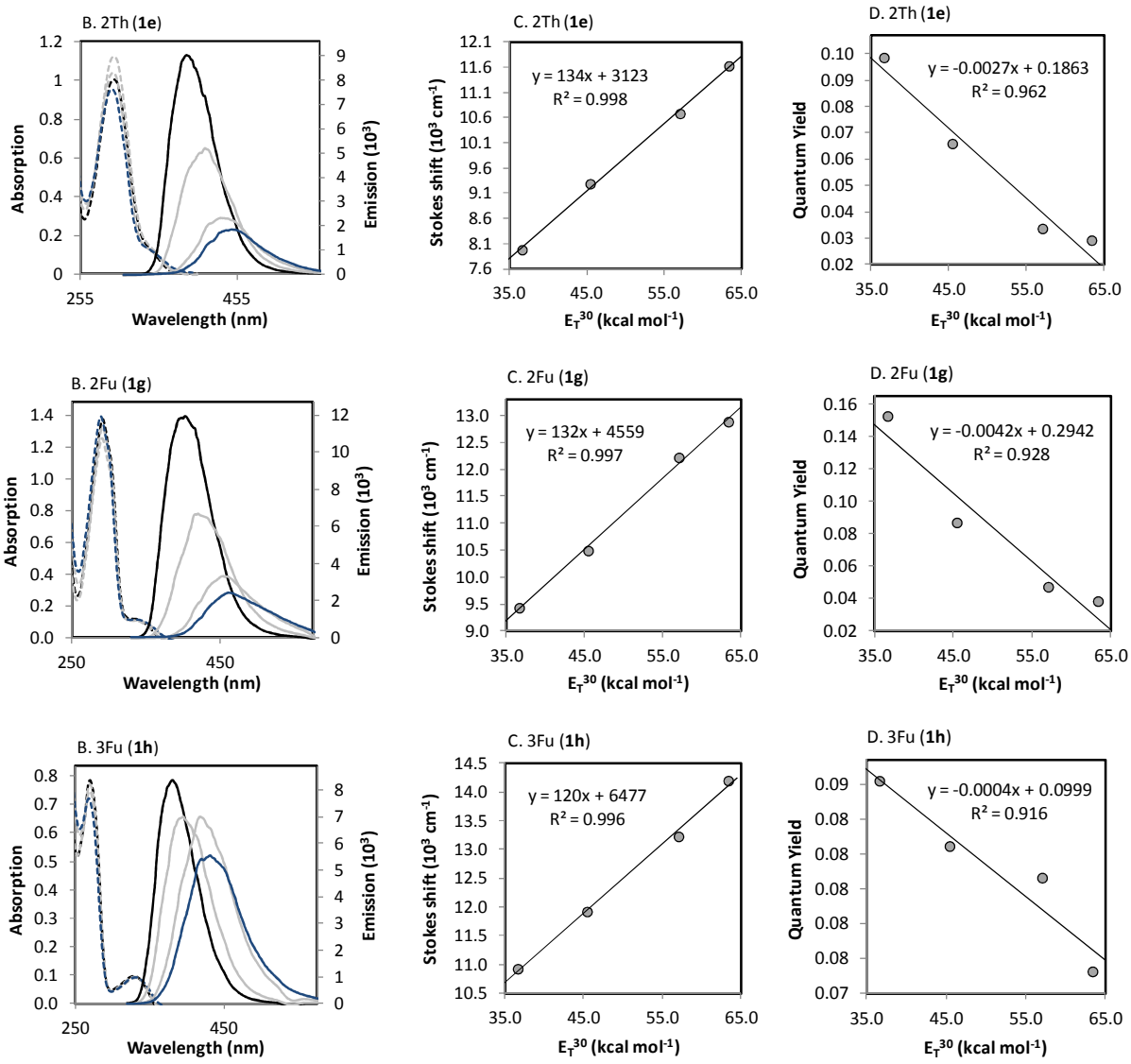


Figure A2. (B) Absorption (dashed) and emission (solid) spectra in water (blue) and dioxane (black) and mixtures (grey) at 40 μM. (C) Linear relationship between stokes shift and solvent polarity of in dioxane, water and mixtures. (D) Linear relationship between quantum yield and solvent polarity of in dioxane, water and mixtures.

A3. Fluorescence Spectra of *N*-Containing Biaryl Series

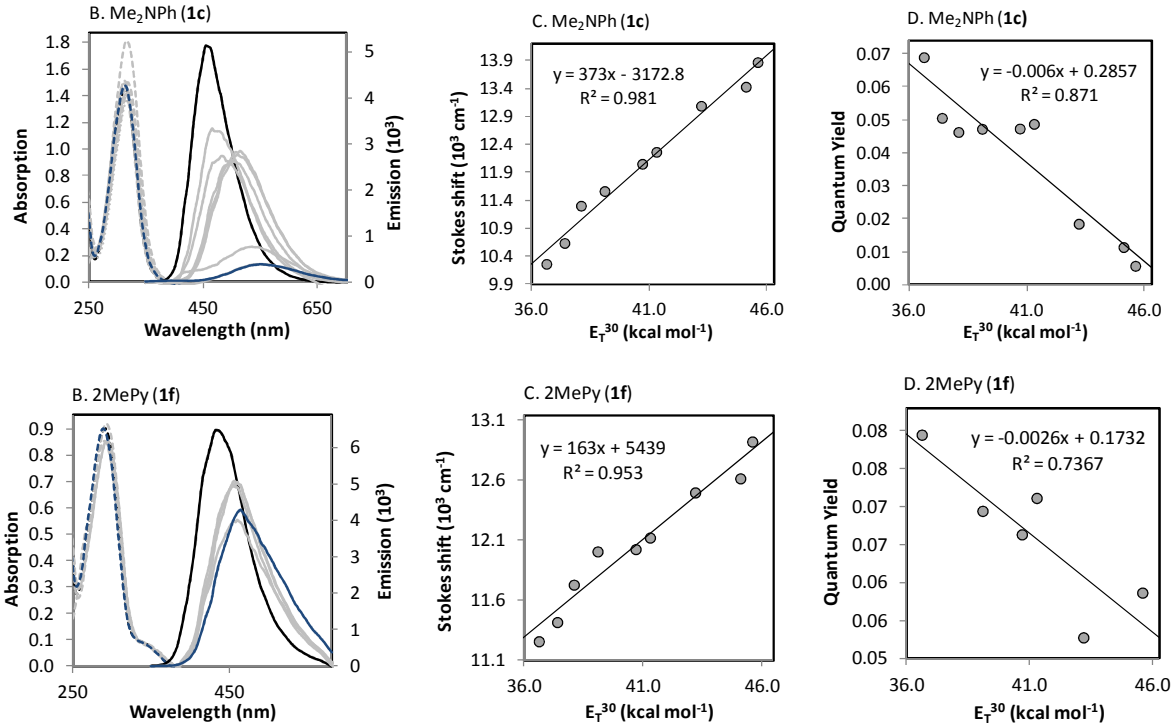
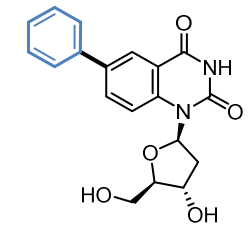
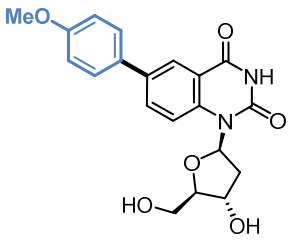


Figure A3. (B) Absorption (dashed) and emission (solid) spectra of **Me₂NPh** and **2MePy** (40 μM) in MeCN (blue) and dioxane (black) and intermediate solvents (grey). (C) Linear relationship between stokes shift and solvent polarity in solvent polarity ranging from dioxane to MeCN. (D) Linear relationship between quantum yield and solvent polarity in solvent polarity ranging from dioxane to MeCN. In order obtain a sensitivity slope to realize a comparison study with the other fluorescent thymidine nucleosides, some data points were omitted during the linearization of **2MePy**.

A4. Fluorescence Data of Six-Membered-Ring Biaryl Series

Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	270	320	371	10.08	0.110
	0.90 - 0.10	226	270	322	382	10.86	0.096
	0.30 - 0.70	222	270	323	395	11.72	0.085
	H_2O	-	269	322	404	12.42	0.085
	D_2O	-	268	324	402	12.44	0.075

Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	280	323	390	10.07	0.132
	0.90 - 0.10	226	280	327	415	11.62	0.087
	0.30 - 0.70	222	280	330	451	13.54	0.074
	H_2O	240	276	330	463	14.63	0.035
	D_2O	240	276	329	462	14.59	0.071

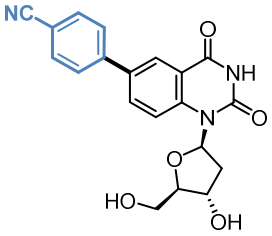
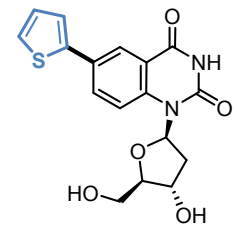
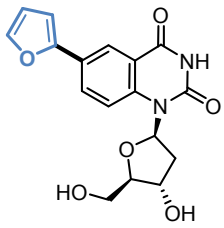
Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	293	362	6.51	0.206
	0.90 - 0.10	226	292	372	7.37	0.197
	0.30 - 0.70	222	289	383	8.49	0.145
	H_2O	245	286	388	9.19	0.140
	D_2O	244	287	388	9.07	0.164

Table A4. Photophysical data of **Ph**, **MeOPh** and **CNPh** in solvent mixtures of dioxane and water. [a] λ_{abs} are reported in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1}. [d] Quantum yield values were measured at λ_{abs} = 270 nm (**Ph**), 280 nm (**MeOPh**) and 290 nm (**CNPh**) using 2-Aminopyridine in 0.1N H_2SO_4 (ϕ = 0.60) as the standard. Errors associated are ± 10% of the given values.

A5. Fluorescence Data of Five-Membered-Ring Biaryl Series

Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	298	391	7.98	0.094
	0.90 - 0.10	226	299	414	9.29	0.061
	0.30 - 0.70	220	297	435	10.68	0.029
	H_2O	-	295	449	11.63	0.024
	D_2O	-	294	447	11.64	0.032

Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	292	403	9.43	0.153
	0.90 - 0.10	226	292	421	10.49	0.087
	0.30 - 0.70	220	292	454	12.22	0.047
	H_2O	220	290	463	12.88	0.038
	D_2O	220	289	463	13.00	0.044

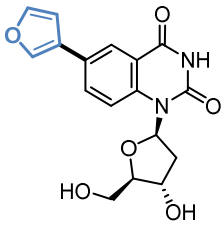
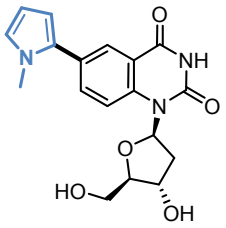
Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	269	330	381	10.93	0.087
	0.90 - 0.10	226	269	330	396	11.92	0.083
	0.30 - 0.70	222	270	332	420	13.23	0.081
	H_2O	228	267	331	430	14.20	0.076
	D_2O	-	267	330	430	14.20	0.089

Table A5. Photophysical data of **2Th**, **2Fu** and **3Fu** in solvent mixtures of dioxane and water. [a] λ_{abs} are reported in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1}. [d] Quantum yield values were measured at λ_{abs} = 298 nm (**2Th**), 292 nm (**2Fu**) and 270 nm (**3Fu**) using 2-Aminopyridine in 0.1N H_2SO_4 (ϕ = 0.60) as the standard. Errors associated are ± 10% of the given values.

A6. Fluorescence Data of *N*-Containing Biaryl Series

Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	291	433	11.27	0.080
	THF	228	291	436	11.43	0.060
	EtOAc	-	291	442	11.74	0.055
	$CHCl_3$	238	295	457	12.02	0.069
	CH_2Cl_2	232	294	455	12.04	0.066
	DCE	228	294	457	12.13	0.071
	DMF	-	292	460	12.51	0.053
	DMSO	-	293	465	12.62	0.076
	MeCN	226	290	464	12.93	0.059
	H_2O	-	286	375	8.30	0.0002
	D_2O	-	286	375	8.30	0.0002

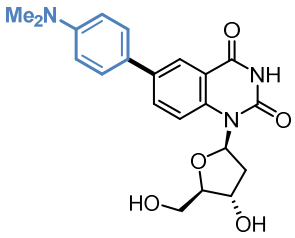
Structure	Solvent	λ_{abs}[a]	λ_{abs}[a]	λ_{em}[b]	Stokes[c]	ϕ[d]
	Dioxane	228	312	459	10.27	0.071
	THF	228	312	467	10.64	0.053
	EtOAc	-	312	482	11.30	0.048
	$CHCl_3$	238	318	503	11.57	0.049
	CH_2Cl_2	230	317	513	12.05	0.049
	DCE	228	316	516	12.27	0.051
	DMF	-	315	536	13.09	0.020
	DMSO	-	316	549	13.43	0.013
	MeCN	226	312	550	13.87	0.008
	H_2O	-	297	460	11.91	0.010
	D_2O	-	296	460	12.05	0.027

Table A6. Photophysical data of **Me2Py** and **Me_2NPh** in organic solvents. [a] λ_{abs} are reported in nm. [b] λ_{em} in nm. [c] Stokes shifts are reported in 10^3 cm^{-1}. [d] Quantum yield values were measured at λ_{abs} = 290 nm (**Me2Py**) and 312 nm (**Me_2NPh**) using 2-Aminopyridine in 0.1N H_2SO_4 (ϕ = 0.60) as the standard. Errors associated are ± 10% of the given values.

A7. Molecular Orbital Plots of the Six-Membered Rings

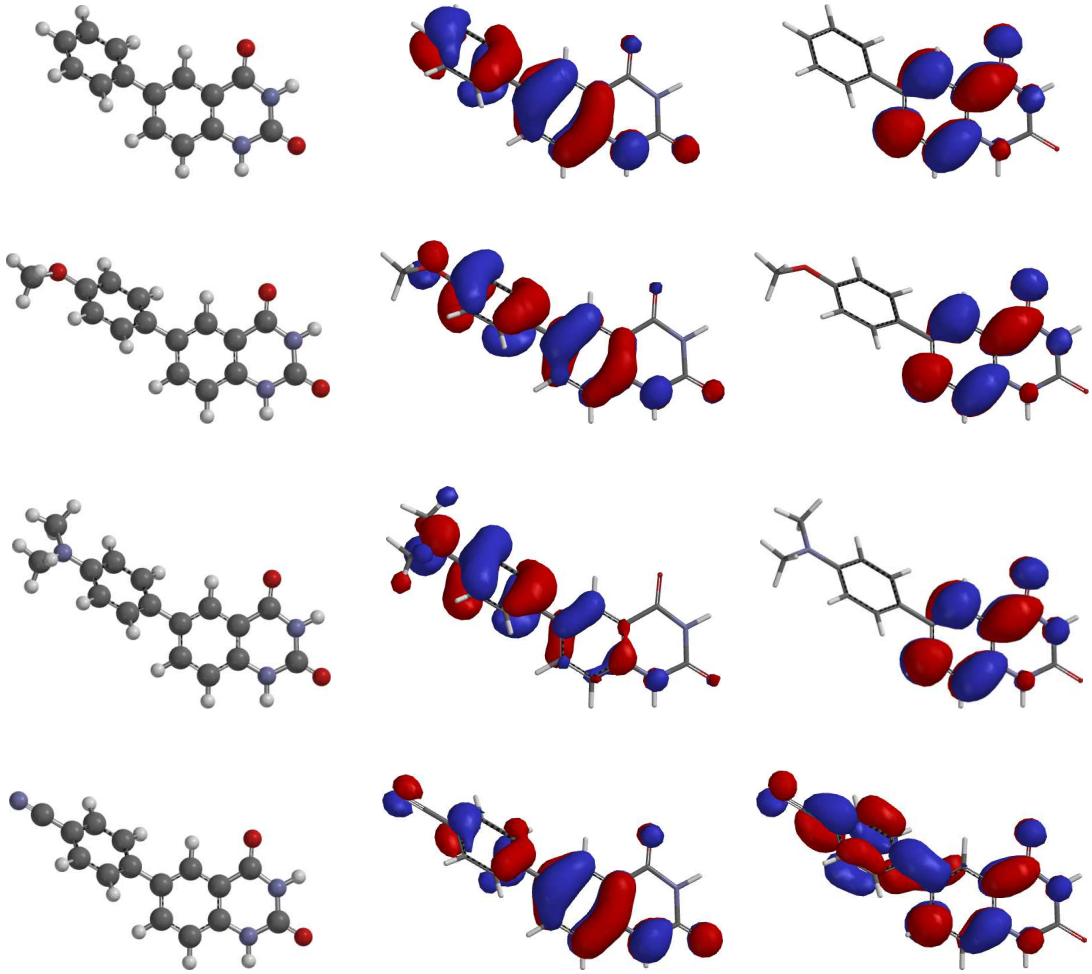


Figure A7. Molecular orbital plots of the isolated nucleobases of **50a – d**. Structure and HOMO / LUMO orbitals from DFT-optimized geometries using B3LYP/6-31G.

A8. Molecular Orbital Plots of the Five-Membered Rings.

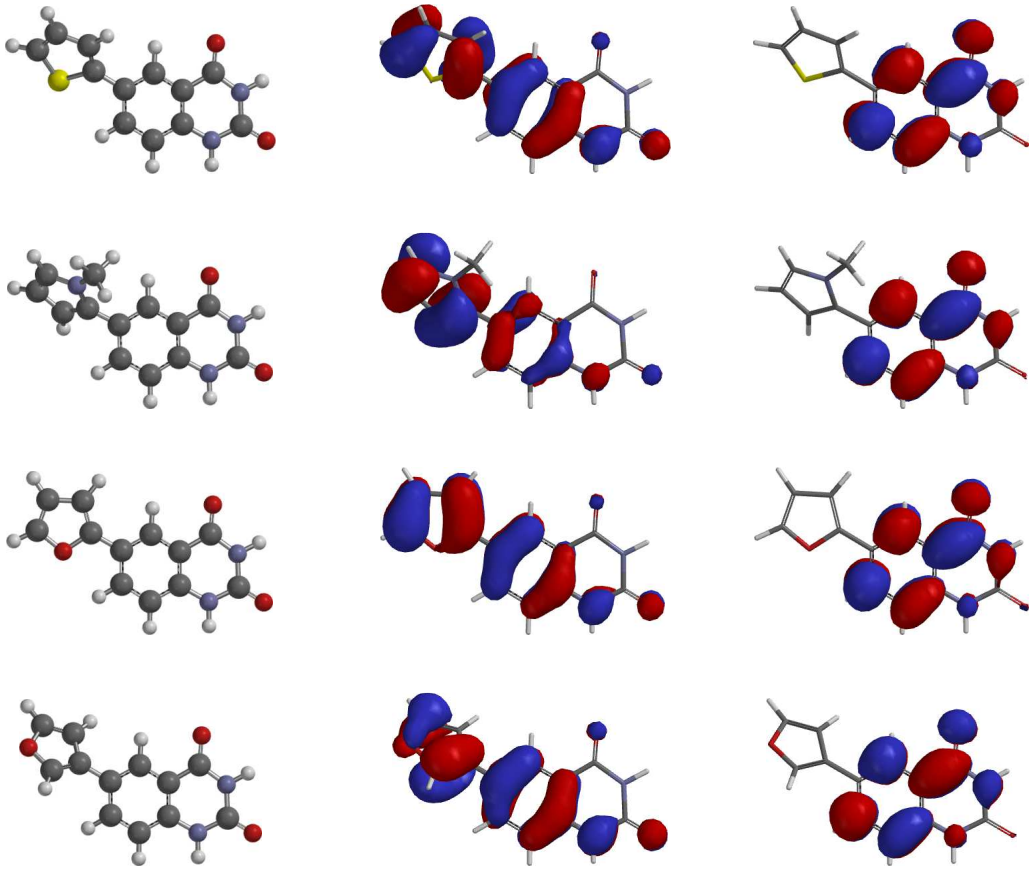


Figure A8. Molecular orbital plots of the isolated nucleobases of **50e – h**. Structure and HOMO / LUMO orbitals from DFT-optimized geometries using B3LYP/6-31G.

A9. Circular Dichroism of *i*-motif DNA

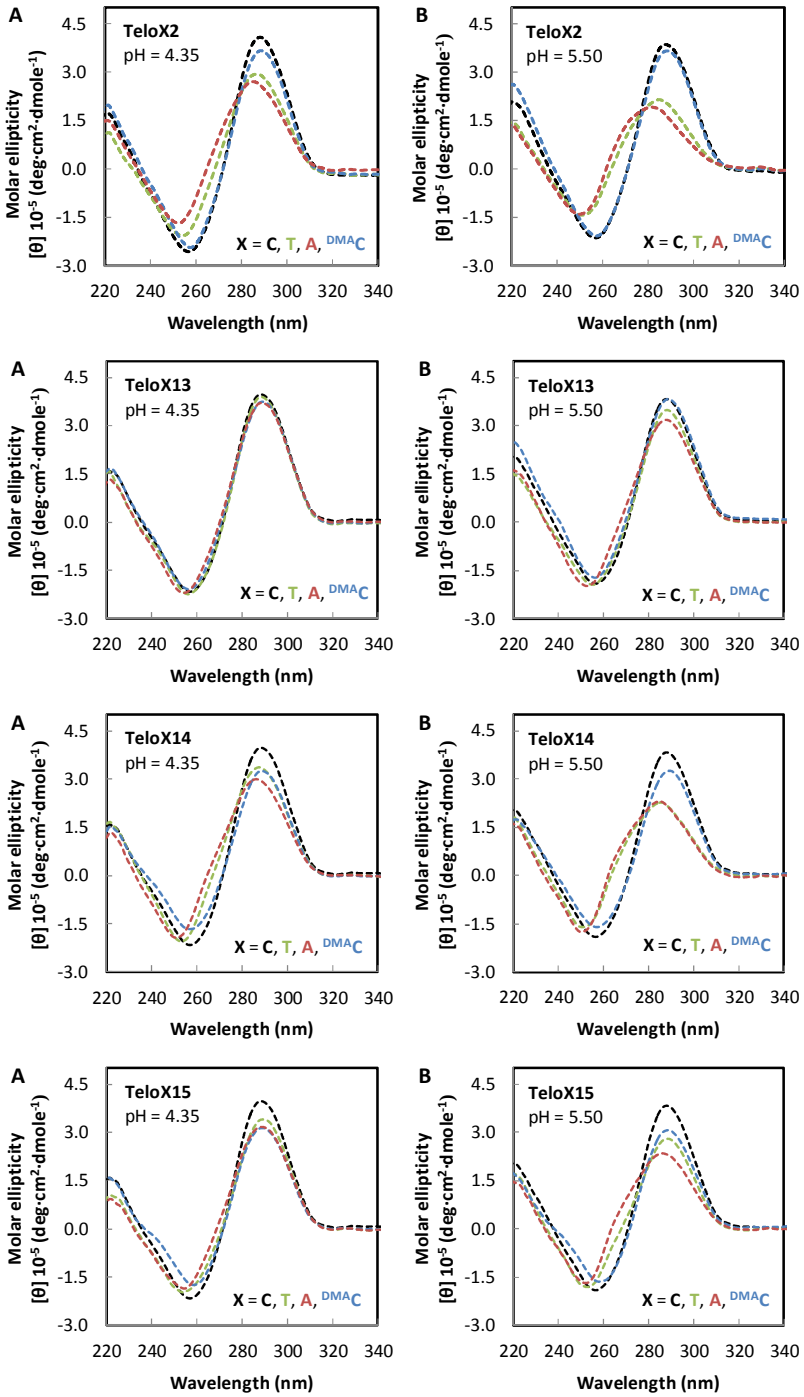


Figure A9. CD spectra at 25 °C of the DNA *i*-motif of **TeloX2**, **X13**, **X14** and **X15** with **X** = **C**, **T**, **A** and ^{DMA}**C** at (A) pH = 4.35 and at (B) pH = 5.50. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl).

A10. Thermal Denaturation of i-motif DNA

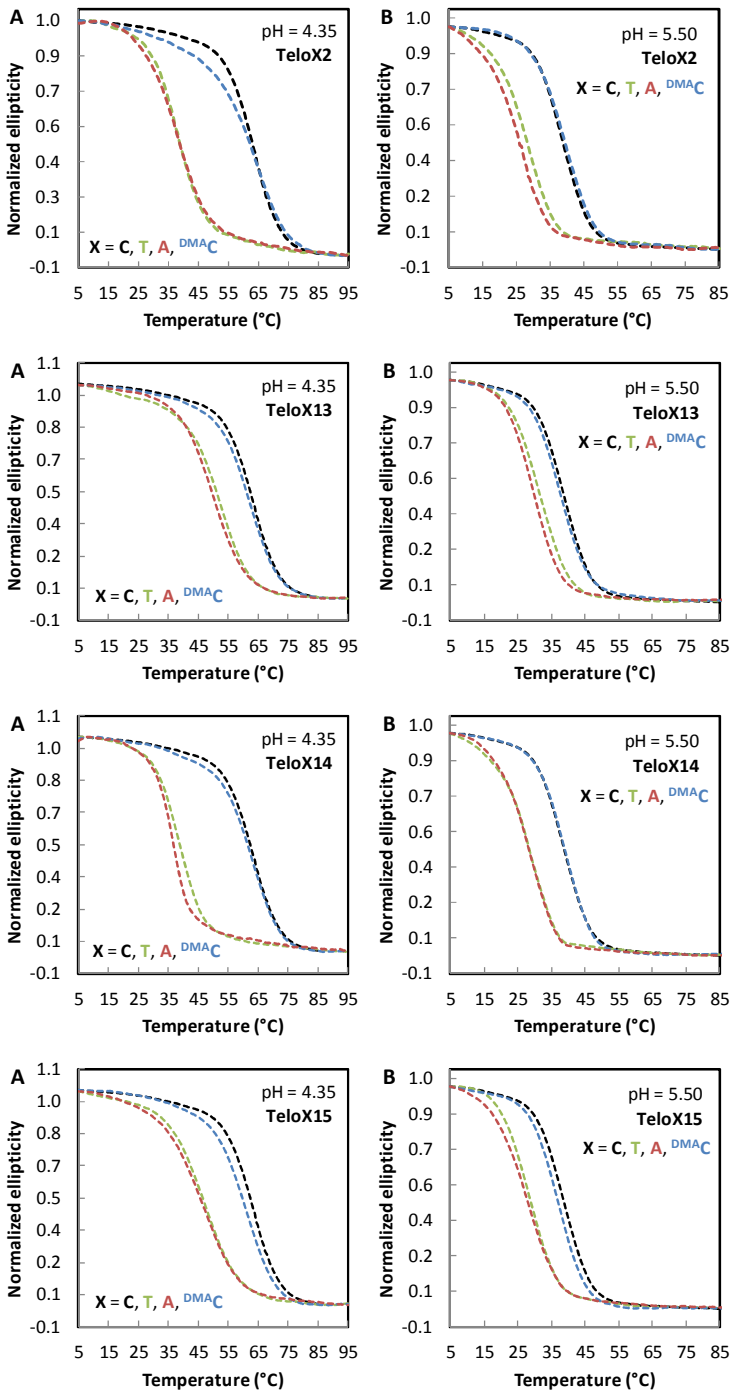


Figure A10. Thermal denaturation spectra of the DNA *i*-motif of **TeloX2**, **X13**, **X14** and **X15** with **X** = **C**, **T**, **A** and ^{DMA}**C** at (A) pH = 4.35 and at (B) pH = 5.50. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl).

A11. Circular Dichroism and Thermal Denaturation of Duplex DNA

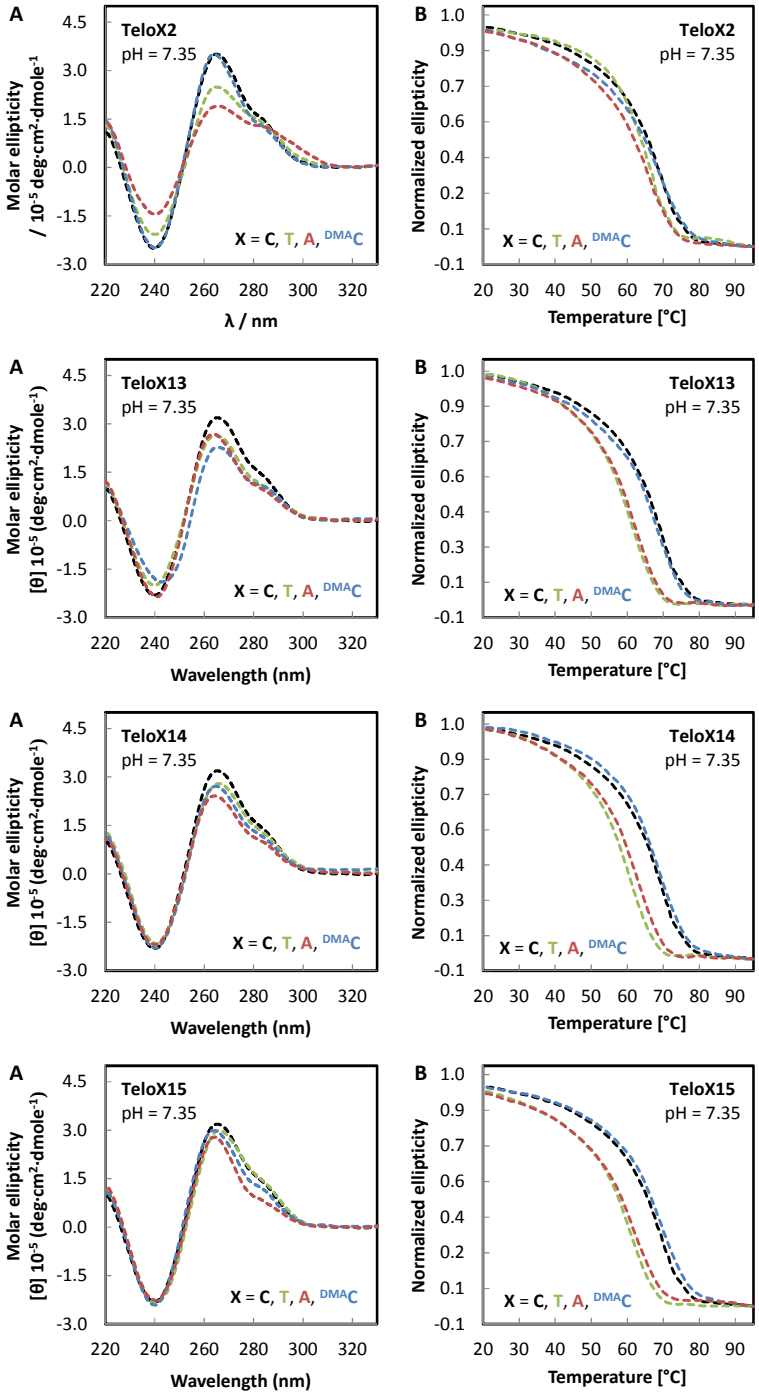


Figure A11. (A) CD at 25 °C and (B) thermal denaturation spectra of DNA duplex of **TeloX2**, **X13**, **X14** and **X15** with **X** = **C**, **T**, **A** and ^{DMA}**C** at pH = 7.35. All samples contained 5 μM of DNA in phosphate citric acid buffer (200 mM of Na_2HPO_4, 100 mM of citric acid and 100 mM NaCl). Fluorescence in DNA

A12. Fluorescence of ^{DMA}C-Modified DNA

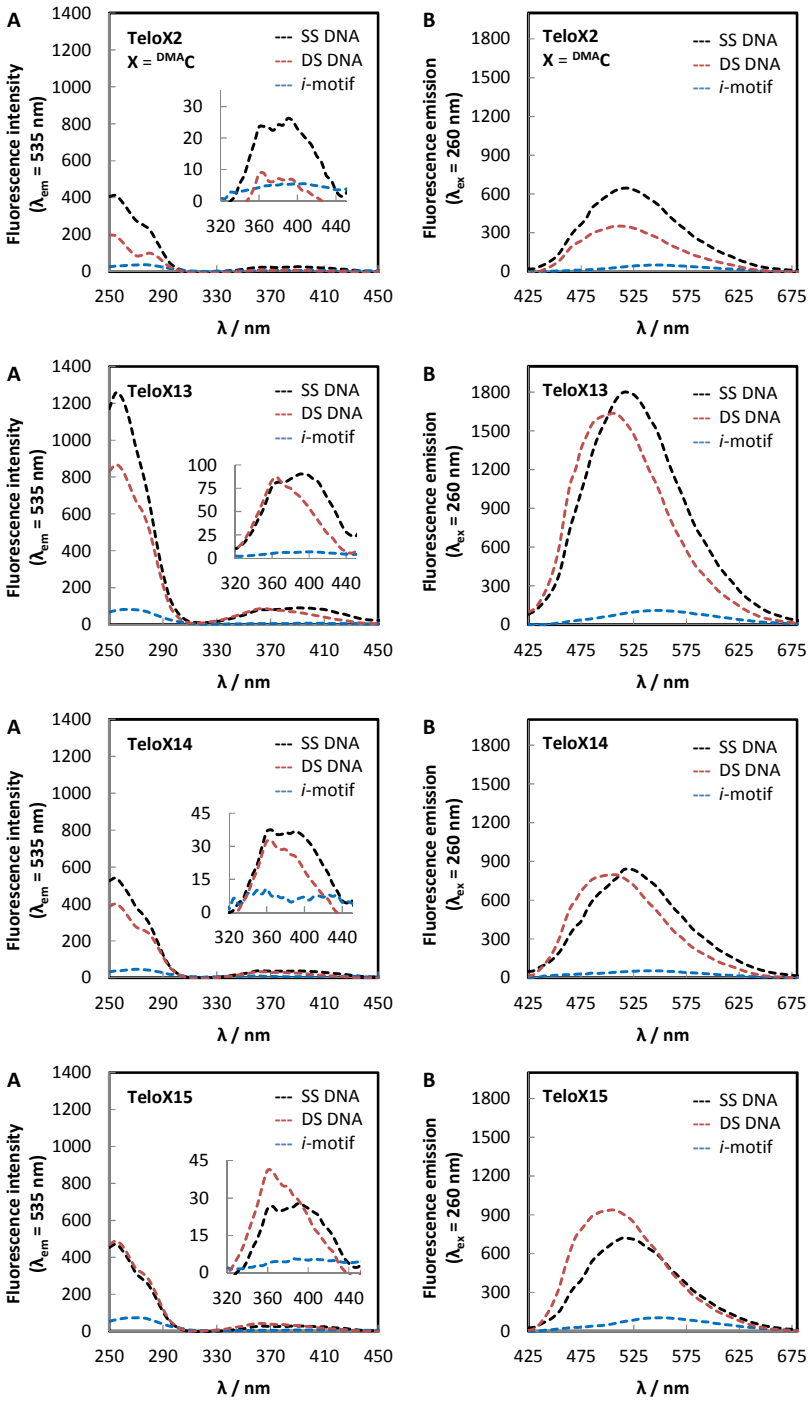


Figure A12. (A) Excitation (λ_{em} = 535 nm) and (B) fluorescence spectra (λ_{ex} = 260 nm) of **TeloX2**, **TeloX13**, **TeloX14** and **TeloX15** with **X** = ^{DMA}**C** as an unfolded (black, SS), duplex (red, DS) and *i*-motif DNA (blue). All samples contained 4 μM DNA in buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl at pH = 7.35).

A13. Strand Displacement for Different "toehold" Lengths at pH = 7.35

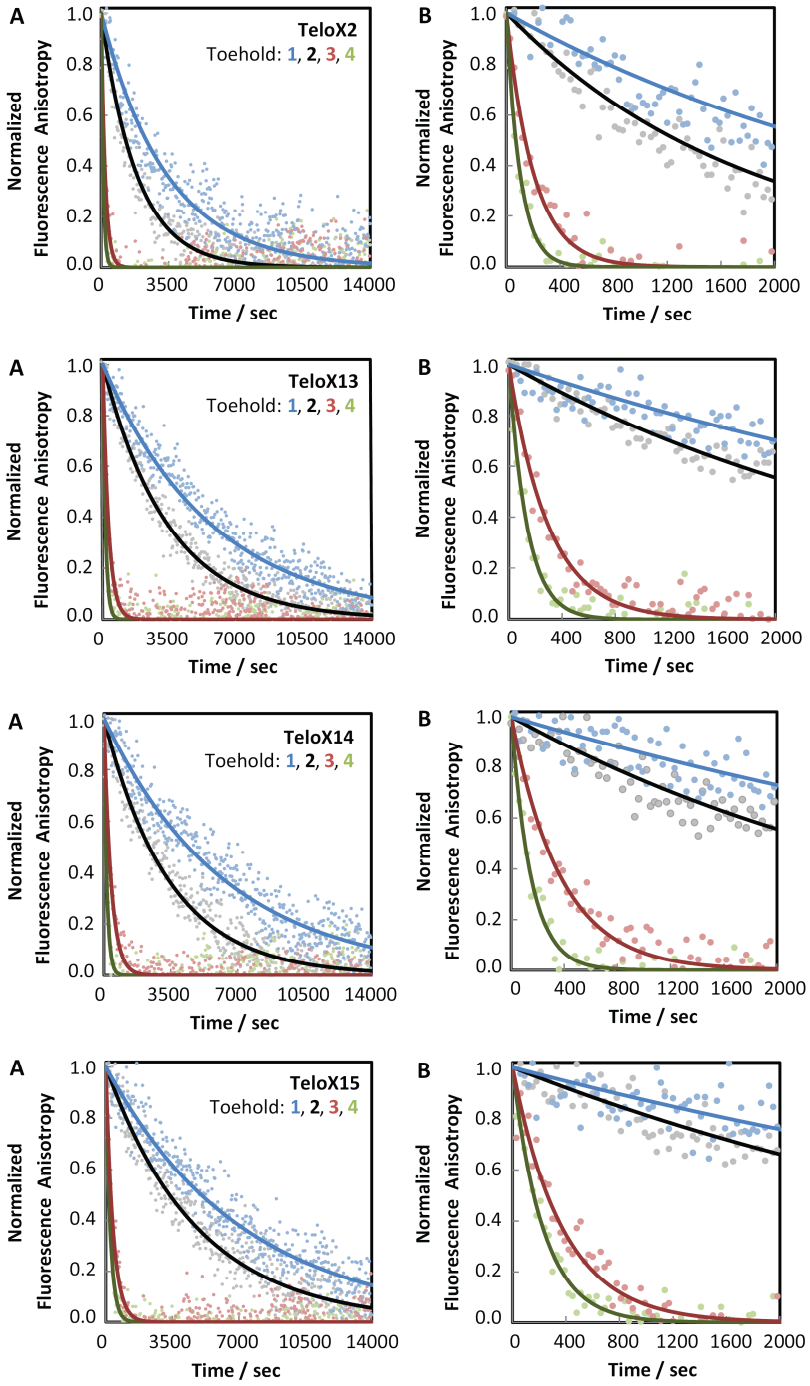


Figure A13. Strand displacement reactions utilizing different toehold lengths according to fluorescence anisotropy of ^{DMA}**C**-modified oligonucleotides (λ_{ex} = 370 nm and λ_{em} = 535 nm) from (A) 0 to 14000 sec; and (B) 0 to 2000 sec (**X** = ^{DMA}**C**). The lines represent mono-exponential fits of the data to eq. 4.

A14. Fluorescence of ^{DMA}T-Modified DNA

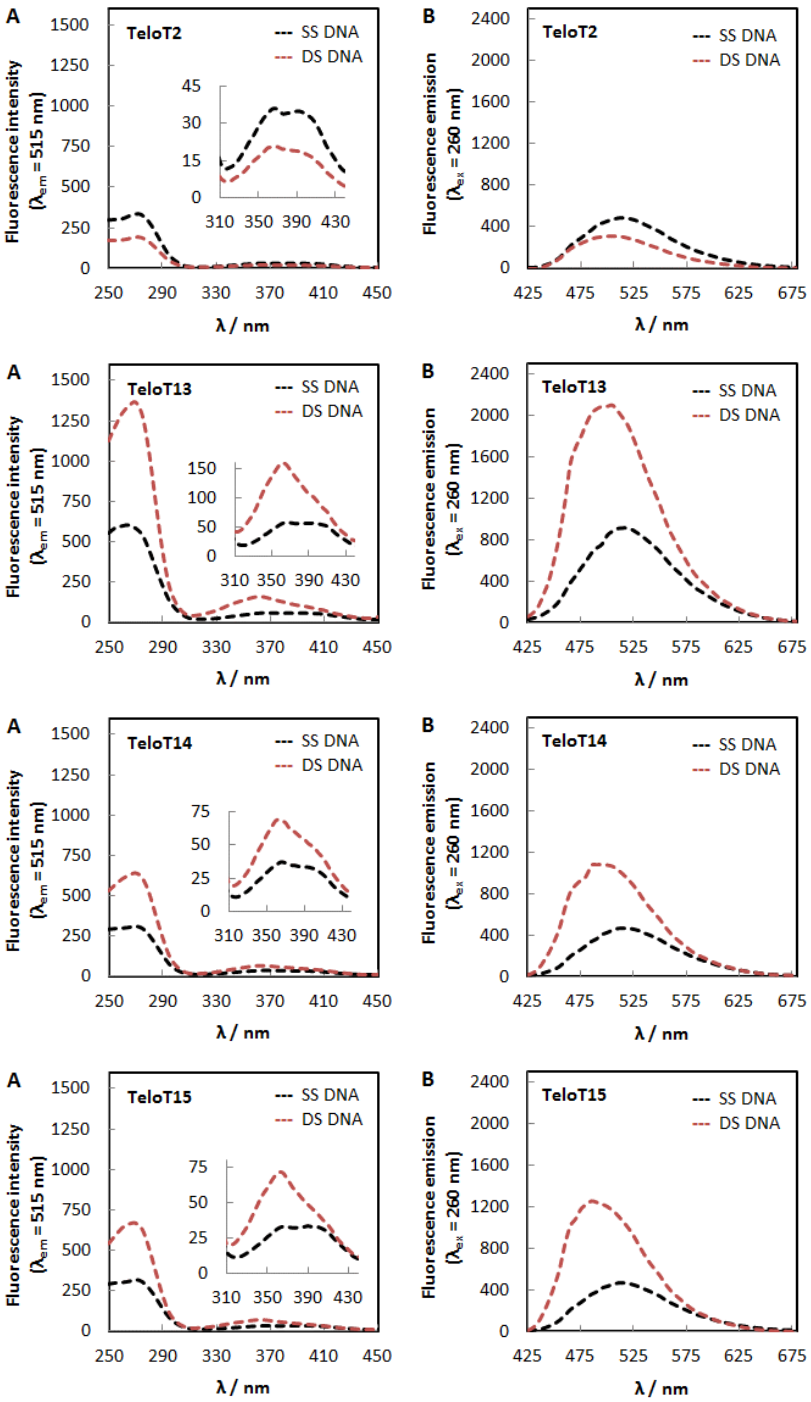


Figure A14. (A) Excitation (λ_{em} = 515 nm) and (B) fluorescence spectra (λ_{ex} = 260 nm) of **TeloX2**, **TeloX13**, **TeloX14** and **TeloX15** with **X** = ^{DMA}**T** as unfolded (black, SS), duplex (red, DS) and *i*-motif DNA (blue). All samples contained 4 μM DNA in buffer (20 mM of Na_2HPO_4, 10 mM of citric acid and 10 mM NaCl at pH = 7.35).

Vita

Academic Profile

Sept. 2010 **Doctoral Studies**

"Fluorescent Nucleosides for Probing Proton-Coupled DNA Folding in Real Time "

University of Zürich, Switzerland, Department of Chemistry, Prof. Dr. Nathan W. Luedtke

Sept. 2008 **Master of Science in Chemistry**

"Design, Synthesis and Analysis of Biologically Active Compounds"

University of Orléans, France, Grade: 16.4/20 with 1st Position.

Sept. 2007 **Bachelor of Science in Chemistry**

University of Orléans, France, Grade: 13.4/20.

Sept. 2005 **Diploma of Higher Education in Chemistry**

University of Strasbourg, France, Grade: 12.6/20.

Scientific Experience

Feb. 2010 - Aug. 2010 **Master Thesis**

"Design, Synthesis and Evaluation of Sequence Specific HIV-1 RNA Ligands as Antiviral Agents"

University of Nice Sophia-Antipolis, France, Laboratoire de Chimie des Molécules Bioactives et des Arômes, Dr. M. Duca and Dr. R. Benhida.

Apr. 2009 - Sept. 2009 **Master Thesis**

"Synthesis of Peptide Conjugates of 5-Aminolaevulinic Acid for Photodynamic Therapy"

University of Bath, United Kingdom, Department of Pharmacy and Pharmacology, Laboratory of Medicinal Chemistry, Dr. I. M. Eggleston.

June. 2009 - Aug. 2009 **Bachelor Thesis**

"Synthesis of Nucleoside Prodrugs as Antiviral Agents"

University of Orléans, France, Department of Organic Chemistry, Institut de Chimie Organique et Analytique, Dr. V. Roy and Prof. Dr. L. Agrofoglio.

Apr. 2007 - Sept. 2007 **Diploma Thesis**

"Solid-Phase Synthesis of Peptides"

NeoMPS pharmaceutical company of Strasbourg, France, which has been recently acquired by PolyPeptide Laboratories. With locations across 3 continents and more than 50 years of experience, the PolyPeptide Group is a world leader in peptide manufacturing.

Awards & Conferences

Sept. 2014 **The SCS Fall Meeting 2014, Switzerland**

Poster: *"Fluorescent Nucleosides for Probing Proton-Coupled DNA Folding in Real Time "*

The Swiss Chemical Society, University of Zürich

June 2014 **The 15th French-American Chemical Society Symposium, France**

Poster: *"Probing DNA Folding Dynamics with Fluorescent Nucleosides"*

Travel Award, the Swiss Academy of Science (SCNAT) and the Swiss Chemical Society

June 2013 **The 43rd National Organic Chemistry Symposium, United States**

Poster: *"Shedding Light on DNA Structure with Fluorescent Tools"*

Travel Award, Graduate School of Chemical and Molecular Sciences Zürich (CMSZH).

Oct. 2012 **The 7th Dorothy Crowfoot Hodgkin Symposium, Switzerland**

Poster: *"Synthesis of Fluorescent Probe for Shedding Light on DNA Structure and Function "*

Best Poster Award, University of Zürich

Sept. 2008 - Sept. 2010 **Master at the University of Orléans, France**

Undergraduate Award: Tow-year merit-based scholarship.

Additional Skills

Languages **French:** Native
English: Fluent
German: Basics (Course at the ETH Zürich)

Teaching **Elaboration and teaching of the exercise lecture** "Organic Chemistry II" (ca. 60 students, 4th semester) in English at the University of Zürich. Supervision of practical courses in general and organic chemistry in English "Organic Chemistry Practicum" (ca. 600 h, 4 semesters). Supervision of undergraduate and master's students.

Publications

7. Article **G. Mata**, N. W. Luedtke, **2015**, *in preparation*.
"A Fluorescent Thymidine Mimic for Probing Mercury Mismatch Pairs in Nucleic Acid."

6. Article **G. Mata**, N. W. Luedtke, *J. Am. Chem. Soc.* **2015**, *137*, 699 - 707.
"Fluorescent Probe for Proton-Coupled DNA Folding Revealing Slow Exchange of i-Motif and Duplex Structures"

5. Article J.P. Joly, **G. Mata**, P. Eldin, L. Briant, F. Fontaine-Vive, M. Duca, R. Benhida, *Chem. Eur. J.* **2014,** *20*, 2071 - 2079.
"Artificial Nucleobase-Amino Acids Conjugates: A new class of TAR RNA Binding Agents"

4. Article **G. Mata**, N. W. Luedtke, *Org. Lett.* **2013**, *15*, 2462 - 2465.
"Synthesis and Solvatochromic Fluorescence of Pyrimidine Biaryl Nucleosides"

3. Article S. M. Huber, **G. Mata**, A. Linden, N. W. Luedtke, *Chem. Commun.* **2013**, *49*, 4280 - 4282.
"Synthesis and Structure of a Hydrogenated Zinc Hemiporphyrazine"

2. Article **G. Mata,** N. W. Luedtke, *J. Org. Chem.* **2012**, *77*, 9006 - 9017.
"Stereoselective N-Glycosylation of 2-Deoxythioribosides for Fluorescent Nucleoside Synthesis"

References

[1] Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561.

[2] Dahm, R. From discovering to understanding Friedrich Miescher's attempts to uncover the function of DNA. *EMBO Rep.* **2010**, *11*, 153.

[3] Chargaff, E.; Magasanik, B.; Doniger, R.; Vischer, E. The Nucleotide Composition of Ribonucleic Acids. *J. Am. Chem. Soc.* **1949**, *71*, 1513.

[4] Todd, A. R. Chemical Structure of the Nucleic Acids. *Proc. Natl. Acad. Sci. U. S. A.* **1954**, *40*, 748.

[5] Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. Molecular Structure of Deoxypentose Nucleic Acids. *Nature* **1953**, *171*, 738.

[6] Franklin, R. E.; Gosling, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, *171*, 740.

[7] Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737.

[8] Watson, J. D.; Crick, F. H. C. Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature* **1953**, *171*, 964.

[9] Hershey, A. D.; Chase, M. Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage. *Journal of General Physiology* **1952**, *36*, 39.

[10] Catasti, P.; Chen, X.; Mariappan, S. V. S.; Bradbury, E. M.; Gupta, G. DNA repeats in the human genome. *Genetica* **1999**, *106*, 15.

[11] Bochman, M. L.; Paeschke, K.; Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **2012**, *13*, 770.

[12] Benabou, S.; Aviñó, A.; Eritja, R.; González, C.; Gargallo, R. Fundamental aspects of the nucleic acid i-motif structures. *RSC Advances* **2014**, *4*, 26956.

[13] Elgar, G.; Vavouri, T. Tuning in to the signals: noncoding sequence conservation in vertebrate genornes. *Trends Genet.* **2008**, *24*, 344.

[14] Ahituv, N.; Zhu, Y.; Visel, A.; Holt, A.; Afzal, V.; Pennacchio, L. A.; Rubin, E. M. Deletion of ultraconserved elements yields viable mice. *PLoS Biol.* **2007**, *5*, 1906.

[15] Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chem.* **2013**, *5*, 182.

[16] Kool, E. T. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1.

[17] Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M. Nucleic Acids in Chemistry and Biology. *Oxford University Press: Oxford, England* **2006**, 1.

[18] Mandal, P. K.; Venkadesh, S.; Gautham, N. Structure of the tetradecanucleotide d(CCCCGGTACCGGGG)(2) as an A-DNA duplex. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **2012**, *68*, 393.

[19] Drew, H. R.; Samson, S.; Dickerson, R. E. Structure of a B-DNA Dodecamer at 16-K. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1982**, *79*, 4040.

[20] Gehring, K.; Leroy, J. L.; Gueron, M. A Tetrameric DNA-Structure with Protonated Cytosine.Cytosine Base-Pairs. *Nature* **1993**, *363*, 561.

[21] Phan, A. T.; Leroy, J. L. Intramolecular i-motif structures of telomeric DNA. *J. Biomol. Struct. Dyn.* **2000**, 245.

[22] Lieblein, A. L.; Buck, J.; Schlepckow, K.; Furtig, B.; Schwalbe, H. Time-Resolved NMR Spectroscopic Studies of DNA i-Motif Folding Reveal Kinetic Partitioning. *Angew. Chem., Int. Ed.* **2012**, *51*, 250.

[23] Lieblein, A. L.; Furtig, B.; Schwalbe, H. Optimizing the Kinetics and Thermodynamics of DNA i-Motif Folding. *ChemBioChem* **2013**, *14*, 1226.

[24] Lieblein, A. L.; Kramer, M.; Dreuw, A.; Furtig, B.; Schwalbe, H. The Nature of Hydrogen Bonds in Cytidine -- H^+ -- Cytidine DNA Base Pairs. *Angew. Chem., Int. Ed.* **2012**, *51*, 4067.

[25] Esmaili, N.; Leroy, J. L. i-motif solution structure and dynamics of the d(AACCCC) and d(CCCCAA) tetrahymena telomeric repeats. *Nucleic Acids Res.* **2005**, *33*, 213.

[26] Kang, C. H.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. Crystal structure of intercalated four-stranded d(C3T) at 1.4 A resolution. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11636.

[27] Chen, L.; Cai, L.; Zhang, X.; Rich, A. Crystal structure of a four-stranded intercalated DNA: d(C4). *Biochemistry* **1994**, *33*, 13540.

[28] Berger, I.; Egli, M.; Rich, A. Inter-strand C-H ··· O hydrogen bonds stabilizing four-stranded intercalated molecules: Stereoelectronic effects of 04' in cytosine-rich DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 12116.

[29] Yang, B.; Rodgers, M. T. Base-Pairing Energies of Proton-Bound Heterodimers of Cytosine and Modified Cytosines: Implications for the Stability of DNA i-Motif Conformations. *J. Am. Chem. Soc.* **2014**, *136*, 282.

[30] Bhaysar-Jog, Y. P.; Van Dornshuld, E.; Brooks, T. A.; Tschumper, G. S.; Wadkins, R. M. Epigenetic Modification, Dehydration, and Molecular Crowding Effects on the Thermodynamics of i-Motif Structure Formation from C-Rich DNA. *Biochemistry* **2014**, *53*, 1586.

[31] Christensen, J. J.; Rytting, J. H.; Izatt, R. M. Thermodynamics of Proton Dissociation in Dilute Aqueous Solution .8. Pk Deliah Degrees and Deltas Degrees Values for Proton Ionization from Several Pyrimidines and Their Nucleosides at 25 Degrees. *J. Phys. Chem.* **1967**, *71*, 2700.

[32] Izatt, R. M.; Christensen, J. J.; Rytting, J. H. Sites and Thermodynamic Quantities Associated with Proton and Metal Ion Interaction with Ribonucleic Acid, Deoxyribonucleic Acid, and Their Constituent Bases, Nucleosides, and Nucleotides. *Chem. Rev.* **1971**, *71*, 439.

[33] Mergny, J. L.; Lacroix, L.; Han, X. G.; Leroy, J. L.; Helene, C. Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an I-DNA Motif. *J. Am. Chem. Soc.* **1995**, *117*, 8887.

[34] Brooks, T. A.; Kendrick, S.; Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* **2010**, *277*, 3459.

[35] Li, X.; Peng, Y. H.; Ren, J. S.; Qu, X. G. Carboxyl-modified single-walled carbon nanotubes selectively induce human telomeric i-motif formation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 19658.

[36] Rajendran, A.; Nakano, S.; Sugimoto, N. Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* **2010**, *46*, 1299.

[37] Miyoshi, D.; Matsumura, S.; Nakano, S.; Sugimoto, N. Duplex dissociation of telomere DNAs induced by molecular crowding. *J. Am. Chem. Soc.* **2004**, *126*, 165.

[38] Brazier, J. A.; Shah, A.; Brown, G. D. I-Motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.* **2012**, *48*, 10739.

[39] Llopis, J.; McCaffery, J. M.; Miyawaki, A.; Farquhar, M. G.; Tsien, R. Y. Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6803.

[40] Webb, B. A.; Chimenti, M.; Jacobson, M. P.; Barber, D. L. Dysregulated pH: a perfect storm for cancer progression. *Nat. Rev. Cancer* **2011**, *11*, 671.

[41] Huppert, J. L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33*, 2908.

[42] Guo, K.; Gokhale, V.; Hurley, L. H.; Sun, D. Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* **2008**, *36*, 4598.

[43] Khan, N.; Avino, A.; Tauler, R.; Gonzalez, C.; Eritja, R.; Gargallo, R. Solution equilibria of the i-motif-forming region upstream of the B-cell lymphoma-2 P1 promoter. *Biochimie* **2007**, *89*, 1562.

[44] Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. *J. Am. Chem. Soc.* **2007**, *129*, 10220.

[45] Sun, D.; Hurley, L. H. The Importance of Negative Superhelicity in Inducing the Formation of G-Quadruplex and i-Motif Structures in the *c-Myc* Promoter: Implications for Drug Targeting and Control of Gene Expression. *J. Med. Chem.* **2009**, *52*, 2863.

[46] Xu, Y.; Sugiyama, H. Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). *Nucleic Acids Res.* **2006**, *34*, 949.

[47] Day, H. A.; Pavlou, P.; Waller, Z. A. E. i-Motif DNA: Structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **2014**, *22*, 4407.

[48] Marsich, E.; Piccini, A.; Xodo, L. E.; Manzini, G. Evidence for a HeLa nuclear protein that binds specifically to the single-stranded d(CCCTAA)(n) telomeric motif. *Nucleic Acids Res.* **1996**, *24*, 4029.

[49] Marsich, E.; Xodo, L. E.; Manzini, G. Widespread presence in mammals and high binding specificity of a nuclear protein that recognises the single-stranded telomeric motif (CCCTAA)(n). *Eur. J. Biochem.* **1998**, *258*, 93.

[50] Kendrick, S.; Kang, H. J.; Alam, M. P.; Madathil, M. M.; Agrawal, P.; Gokhale, V.; Yang, D. Z.; Hecht, S. M.; Hurley, L. H. The Dynamic Character of the BCL2 Promoter i-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J. Am. Chem. Soc.* **2014**, *136*, 4161.

[51] Kang, H. J.; Kendrick, S.; Hecht, S. M.; Hurley, L. H. The Transcriptional Complex Between the BCL2 i-Motif and hnRNP LL Is a Molecular Switch for Control of Gene Expression That Can Be Modulated by Small Molecules. *J. Am. Chem. Soc.* **2014**, *136*, 4172.

[52] Cimino-Reale, G.; Pascale, E.; Alvino, E.; Starace, G.; D'Ambrosio, E. Long telomeric C-rich 5 '-tails in human replicating cells. *J. Biol. Chem.* **2003**, *278*, 2136.

[53] Stewart, S. A.; Weinberg, R. A. Senescence: does it all happen at the ends? *Oncogene* **2002**, *21*, 627.

[54] Blackburn, E. H. Switching and signaling at the telomere. *Cell* **2001**, *106*, 661.

[55] Reed, J. E.; Arnal, A. A.; Neidle, S.; Vilar, R. Stabilization of G-quadruplex DNA and inhibition of telomerase activity by square-planar nickel(II) complexes. *J. Am. Chem. Soc.* **2006**, *128*, 5992.

[56] Sun, D. D.; Liu, Y. A.; Liu, D.; Zhang, R.; Yang, X. C.; Liu, J. Stabilization of G-Quadruplex DNA, Inhibition of Telomerase Activity and Live Cell Imaging Studies of Chiral Ruthenium(II) Complexes. *Chem.--Eur. J.* **2012**, *18*, 4285.

[57] Han, H. Y.; Hurley, L. H. G-quadruplex DNA: a potential target for anti-cancer drug design. *Trends Pharmacol. Sci.* **2000**, *21*, 136.

[58] Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. *Science* **1994**, *266*, 2011.

[59] Phan, A. T.; Mergny, J. L. Human telomeric DNA: G-quadruplex, i-motif and watson-crick double helix. *Nucleic Acids Res.* **2002**, *30*, 4618.

[60] Sun, D. Y.; Guo, K. X.; Rusche, J. J.; Hurley, L. H. Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res.* **2005**, *33*, 6070.

[61] Baldrich, E.; O'Sullivan, C. K. Ability of thrombin to act as molecular chaperone, inducing formation of quadruplex structure of thrombin-binding aptamer. *Anal. Biochem.* **2005**, *341*, 194.

[62] Kouzine, F.; Sanford, S.; Elisha-Feil, Z.; Levens, D. The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.* **2008**, *15*, 146.

[63] Sen, D.; Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **1988**, *334*, 364.

[64] Sundquist, W. I.; Klug, A. Telomeric DNA Dimerizes by Formation of Guanine Tetrads between Hairpin Loops. *Nature* **1989**, *342*, 825.

[65] Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402.

[66] Marusic, M.; Sket, P.; Bauer, L.; Viglasky, V.; Plavec, J. Solution-state structure of an intramolecular G-quadruplex with propeller, diagonal and edgewise loops. *Nucleic Acids Res.* **2012**, *40*, 6946.

[67] Hud, N. V.; Plavec, J. The Role of Cations in Determining Quadruplex Structure and Stability. *Quadruplex Nucleic Acids* **2006**, 100.

[68] Miyoshi, D.; Fujimoto, T.; Sugimoto, N. Molecular Crowding and Hydration Regulating of G-Quadruplex Formation. *Quadruplex Nucleic Acids* **2013**, *330*, 87.

[69] Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix Formation by Guanylic Acid. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013.

[70] Zhang, A. Y. Q.; Balasubramanian, S. The Kinetics and Folding Pathways of Intramolecular G-Quadruplex Nucleic Acids. *J. Am. Chem. Soc.* **2012**, *134*, 19297.

[71] Guedin, A.; Alberti, P.; Mergny, J. L. Stability of intramolecular quadruplexes: sequence effects in the central loop. *Nucleic Acids Res.* **2009**, *37*, 5559.

[72] Gepshtein, R.; Huppert, D.; Lubitz, I.; Amdursky, N.; Kotlyar, A. B. Radiationless transitions of G4 wires and dGMP. *J. Phys. Chem. C* **2008**, *112*, 12249.

[73] Miannay, F. A.; Banyasz, A.; Gustavsson, T.; Markovitsi, D. Excited States and Energy Transfer in G-Quadruplexes. *J. Phys. Chem. C* **2009**, *113*, 11760.

[74] Luedtke, N. W.; Dumas, A. Cation-Mediated Energy Transfer in G-Quadruplexes Revealed by an Internal Fluorescent Probe. *J. Am. Chem. Soc.* **2010**, *132*, 18004.

[75] Sproviero, M.; Fadock, K. L.; Witham, A. A.; Manderville, R. A.; Sharma, P.; Wetmore, S. D. Electronic tuning of fluorescent 8-aryl-guanine probes for monitoring DNA duplex-quadruplex exchange. *Chem. Sci.* **2014**, *5*, 788.

[76] Todd, A. K.; Johnston, M.; Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **2005**, *33*, 2901.

[77] Rodriguez, R.; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **2012**, *8*, 301.

[78] Cahoon, L. A.; Seifert, H. S. An Alternative DNA Structure Is Necessary for Pilin Antigenic Variation in Neisseria gonorrhoeae. *Science* **2009**, *325*, 764.

[79] Neidle, S.; Parkinson, G. Telomere maintenance as a target for anticancer drug discovery. *Nat. Rev. Drug Discovery* **2002**, *1*, 383.

[80] Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Boussin, F. D. Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. *Oncogene* **2005**, *24*, 2917.

[81] Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discovery* **2011**, *10*, 261.

[82] Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593.

[83] Katz, S. Reversible Reaction of Double-Stranded Polynucleotides and Hgii - Separation of Strands. *Nature* **1962**, *195*, 997.

[84] Katz, S. Reversible Reaction of Hg (Ii) and Double-Stranded Polynucleotides a Step-Function Theory and Its Significance. *Biochim. Biophys. Acta* **1963**, *68*, 240.

[85] Froystein, N. A.; Sletten, E. Interaction of Mercury(II) with the DNA Dodecamer CGCGAATTCGCG - a ^{1}H and ^{15}N NMR Study. *J. Am. Chem. Soc.* **1994**, *116*, 3240.

[86] Miyake, Y.; Togashi, H.; Tashiro, M.; Yamaguchi, H.; Oda, S.; Kudo, M.; Tanaka, Y.; Kondo, Y.; Sawa, R.; Fujimoto, T.; Machinami, T.; Ono, A. Mercury^{II}-mediated formation of thymine-Hg^{II}-thymine base pairs in DNA duplexes. *J. Am. Chem. Soc.* **2006**, *128*, 2172.

[87] Torigoe, H.; Miyakawa, Y.; Ono, A.; Kozasa, T. Positive cooperativity of the specific binding between Hg^{2+} ion and T:T mismatched base pairs in duplex DNA. *Thermochim. Acta* **2012**, *532*, 28.

[88] Tanaka, Y.; Oda, S.; Yamaguchi, H.; Kondo, Y.; Kojima, C.; Ono, A. ^{15}N-^{15}N J-coupling across Hg^{II}: Direct observation of Hg^{II}-mediated T-T base pairs in a DNA duplex. *J. Am. Chem. Soc.* **2007**, *129*, 244.

[89] Torigoe, H.; Ono, A.; Kozasa, T. Hg^{II} Ion Specifically Binds with T:T Mismatched Base Pair in Duplex DNA. *Chem.--Eur. J.* **2010**, *16*, 13218.

[90] Kondo, J.; Yamada, T.; Hirose, C.; Okamoto, I.; Tanaka, Y.; Ono, A. Crystal Structure of Metallo DNA Duplex Containing Consecutive Watson-Crick-like T-Hg^{II}-T Base Pairs. *Angew. Chem., Int. Ed.* **2014**, *53*, 2385.

[91] Yamaguchi, H.; Sebera, J.; Kondo, J.; Oda, S.; Komuro, T.; Kawamura, T.; Dairaku, T.; Kondo, Y.; Okamoto, I.; Ono, A.; Burda, J. V.; Kojima, C.; Sychrovsky, V.; Tanaka, Y. The structure of

metallo-DNA with consecutive thymine-Hg^{II}-thymine base pairs explains positive entropy for the metallo base pair formation. *Nucleic Acids Res.* **2014**, *42*, 4094.

[92] Pyykko, P. Strong closed-shell interactions in inorganic chemistry. *Chem. Rev.* **1997**, *97*, 597.

[93] Urata, H.; Yamaguchi, E.; Funai, T.; Matsumura, Y.; Wada, S. Incorporation of Thymine Nucleotides by DNA Polymerases through T-Hg^{II}-T Base Pairing. *Angew. Chem., Int. Ed.* **2010**, *49*, 6516.

[94] Horton, N. C.; Perona, J. J. Making the most of metal ions. *Nature Structural Biology* **2001**, *8*, 290.

[95] Kypr, J.; Kejnovska, I.; Renciuk, D.; Vorlickova, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37*, 1713.

[96] Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.* **2005**, *33*.

[97] Dai, J. X.; Ambrus, A.; Hurley, L. H.; Yang, D. Z. A Direct and Nondestructive Approach To Determine the Folding Structure of the I-Motif DNA Secondary Structure by NMR. *J. Am. Chem. Soc.* **2009**, *131*, 6102.

[98] Parkinson, G. N.; Lee, M. P. H.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876.

[99] Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Springer: New York, **2006**.

[100] Neef, A. B.; Luedtke, N. W. Dynamic metabolic labeling of DNA in vivo with arabinosyl nucleosides. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 20404.

[101] Rieder, U.; Luedtke, N. W. Alkene-Tetrazine Ligation for Imaging Cellular DNA. *Angew. Chem.* **2014**.

[102] Neef, A. B.; Luedtke, N. W. An Azide-Modified Nucleoside for Metabolic Labeling of DNA. *ChemBioChem* **2014**, *15*, 789.

[103] Neef, A. B.; Samain, F.; Luedtke, N. W. Metabolic Labeling of DNA by Purine Analogues *in Vivo*. *ChemBioChem* **2012**, *13*, 1750.

[104] Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. Creating new fluorescent probes for cell biology. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906.

[105] Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. Review - The fluorescent toolbox for assessing protein location and function. *Science* **2006**, *312*, 217.

[106] Choi, J.; Kim, S.; Tachikawa, T.; Fujitsuka, M.; Majima, T. pH-Induced Intramolecular Folding Dynamics of i-Motif DNA. *J. Am. Chem. Soc.* **2011**, *133*, 16146.

[107] Liu, D. S.; Balasubramanian, S. A proton-fuelled DNA nanomachine. *Angew. Chem., Int. Ed.* **2003**, *42*, 5734.

[108] Gradinaru, C. C.; Marushchak, D. O.; Samim, M.; Krull, U. J. Fluorescence anisotropy: from single molecules to live cells. *Analyst* **2010**, *135*, 452.

[109] Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **2010**, *110*, 2579.

[110] Daniels, M.; Hauswirth, W. Fluorescence of the purine and pyrimidine bases of the nucleic acids in neutral aqueous solution at 300 degrees K. *Science* **1971**, *171*, 675.

[111] Callis, P. R. Electronic States and Luminescence of Nucleic-Acid Systems. *Annu. Rev. Phys. Chem.* **1983**, *34*, 329.

[112] Okamoto, A.; Tainaka, K.; Nishiza, K.; Saito, I. Monitoring DNA structures by dual fluorescence of pyrene derivatives. *J. Am. Chem. Soc.* **2005**, *127*, 13128.

[113] Okamoto, A.; Tanaka, K.; Fukuta, T.; Saito, I. Design of base-discriminating fluorescent nucleoside and its application to T/C SNP typing. *J. Am. Chem. Soc.* **2003**, *125*, 9296.

[114] Xie, Y.; Dix, A. V.; Tor, Y. FRET Enabled Real Time Detection of RNA-Small Molecule Binding. *J. Am. Chem. Soc.* **2009**, *131*, 17605.

[115] Miyata, K.; Tamamushi, R.; Ohkubo, A.; Taguchi, H.; Seio, K.; Santa, T.; Sekine, M. Synthesis and properties of a new fluorescent bicyclic 4-N-carbamoyldeoxycytidine derivative. *Org. Lett.* **2006**, *8*, 1545.

[116] Greco, N. J.; Sinkeldam, R. W.; Tor, Y. An Emissive C Analog Distinguishes between G, 8-oxoG, and T. *Org. Lett.* **2009**, *11*, 1115.

[117] Srivatsan, S. G.; Greco, N. J.; Tor, Y. A highly emissive fluorescent nucleoside that signals the activity of toxic ribosome-inactivating proteins. *Angewandte Chemie-International Edition* **2008**, *47*, 6661.

[118] Cosa, G.; Focsaneanu, K. S.; McLean, J. R. N.; McNamee, J. P.; Scaiano, J. C. Photophysical properties of fluorescent DNA-dyes bound to single- and double-stranded DNA in aqueous buffered solution. *Photochem. Photobiol.* **2001**, *73*, 585.

[119] Kapuscinski, J.; Darzynkiewicz, Z. Interactions of Acridine-Orange with Double-Stranded Nucleic-Acids - Spectral and Affinity Studies. *J. Biomol. Struct. Dyn.* **1987**, *5*, 127.

[120] Kapuscinski, J. Interactions of Nucleic-Acids with Fluorescent Dyes - Spectral Properties of Condensed Complexes. *J. Histochem. Cytochem.* **1990**, *38*, 1323.

[121] Vummidi, B. R.; Alzeer, J.; Luedtke, N. W. Fluorescent Probes for G-Quadruplex Structures. *ChemBioChem* **2013**, *14*, 540.

[122] Alzeer, J.; Vummidi, B. R.; Roth, P. J. C.; Luedtke, N. W. Guanidinium-Modified Phthalocyanines as High-Affinity G-Quadruplex Fluorescent Probes and Transcriptional Regulators. *Angew. Chem., Int. Ed.* **2009**, *48*, 9362.

[123] Alzeer, J.; Luedtke, N. W. pH-Mediated Fluorescence and G-Quadruplex Binding of Amido Phthalocyanines. *Biochemistry* **2010**, *49*, 4339.

[124] Moreira, B. G.; You, Y.; Behlke, M. A.; Owczarzy, R. Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability. *Biochem. Biophys. Res. Commun.* **2005**, *327*, 473.

[125] Mergny, J. L.; Maurizot, J. C. Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *ChemBioChem* **2001**, *2*, 124.

[126] Dodd, D. W.; Hudson, R. H. E. Intrinsically Fluorescent Base-Discriminating Nucleoside Analogs. *Mini-Rev. Org. Chem.* **2009**, *6*, 378.

[127] Godde, F.; Toulme, J. J.; Moreau, S. 4-amino-1H-benzo[g]quinazoline-2-one: a fluorescent analog of cytosine to probe protonation sites in triplex forming oligonucleotides. *Nucleic Acids Res.* **2000**, *28*, 2977.

[128] Wilhelmsson, L. M.; Holmen, A.; Lincoln, P.; Nielson, P. E.; Norden, B. A highly fluorescent DNA base analogue that forms Watson-Crick base pairs with guanine. *J. Am. Chem. Soc.* **2001**, *123*, 2434.

[129] Greco, N. J.; Tor, Y. Simple fluorescent pyrimidine analogues detect the presence of DNA abasic sites. *J. Am. Chem. Soc.* **2005**, *127*, 10784.

[130] Xie, Y.; Maxson, T.; Tor, Y. Fluorescent Ribonucleoside as a FRET Acceptor for Tryptophan in Native Proteins. *J. Am. Chem. Soc.* **2010**, *132*, 11896.

[131] Okamoto, A.; Kanatani, K.; Saito, I. Pyrene-labeled base-discriminating fluorescent DNA probes for homogeneous SNP typing. *J. Am. Chem. Soc.* **2004**, *126*, 4820.

[132] Ehrenschwender, T.; Wagenknecht, H. A. Synthesis and Spectroscopic Characterization of BODIPY-Modified Uridines as Potential Fluorescent Probes for Nucleic Acids. *Synthesis* **2008**, 3657.

[133] Dumas, A.; Luedtke, N. W. Fluorescence Properties of 8-(2-Pyridyl)guanine "2PyG" as Compared to 2-Aminopurine in DNA. *ChemBioChem* **2011**, *12*, 2044.

[134] Zhang, L.-H.; Xi, Z.; Chattopadhyaya, J. *Medicinal Chemistry of Nucleic Acids*; John Wiley & Sons, Inc., Hoboken, NJ, USA., 2011.

[135] Vittori, S.; Dal Ben, D.; Lambertucci, C.; Marucci, G.; Volpini, R.; Cristalli, G. Antiviral properties of deazaadenine nucleoside derivatives. *Curr. Med. Chem.* **2006**, *13*, 3529.

[136] De Clercq, E. The Discovery of Antiviral Agents: Ten Different Compounds, Ten Different Stories. *Med. Res. Rev.* **2008**, *28*, 929.

[137] Uhlmann, E.; Peyman, A. Antisense Oligonucleotides - a New Therapeutic Principle. *Chem. Rev.* **1990**, *90*, 543.

[138] Englisch, U.; Gauss, D. H. Chemically Modified Oligonucleotides as Probes and Inhibitors. *Angew. Chem., Int. Ed.* **1991**, *30*, 613.

[139] Cobb, A. J. A. Recent highlights in modified oligonucleotide chemistry. *Org. Biomol. Chem.* **2007**, *5*, 3260.

[140] Salic, A.; Mitchison, T. J. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 2415.

[141] Chien, T. C.; Chen, C. S.; Chern, J. W. Nucleosides XIII. Facile synthesis of 4-amino-1-(2-deoxy-beta-D-ribofuranosyl)quinazolin-2-one as a 2 '-deoxycytidine analog for oligonucleotide synthesis. *J. Chin. Chem. Soc.* **2005**, *52*, 1237.

[142] Fossey, C.; Landelle, H.; Laduree, D.; Robba, M. Synthesis of N-1-Beta-D-Arabinofuranosyl and N-1-2'-Deoxy-Beta-D-Erythro-Pentofuranosyl Thieno[3,2-D]Pyrimidine Nucleosides. *Nucleosides Nucleotides* **1994**, *13*, 925.

[143] Tor, Y.; Del Valle, S.; Jaramillo, D.; Srivatsan, S. G.; Rios, A.; Weizman, H. Designing new isomorphic fluorescent nucleobase analogues: the thieno[3,2-d]pyrimidine core. *Tetrahedron* **2007**, *63*, 3608.

[144] Liu, H. B.; Gao, J. M.; Kool, E. T. Size-expanded analogues of dG and dC: Synthesis and pairing properties in DNA. *J. Org. Chem.* **2005**, *70*, 639.

[145] Robins, M. J.; Wilson, J. S. Nucleic-Acid Related-Compounds .32. Smooth and Efficient Deoxygenation of Secondary Alcohols - a General Procedure for the Conversion of Ribonucleosides to 2'-Deoxynucleosides. *J. Am. Chem. Soc.* **1981**, *103*, 932.

[146] Seela, F.; Debelak, H. 8-aza-7-deazaadenine and 7-deazaguanine: Synthesis and properties of nucleosides and oligonucleotides with nucleobases linked at position-8. *Nucleosides Nucleotides & Nucleic Acids* **2001**, *20*, 577.

[147] Park, M.; Rizzo, C. J. Stereocontrolled de Novo synthesis of beta-2'-deoxyribonucleosides. *J. Org. Chem.* **1996**, *61*, 6092.

[148] Wang, Z. W.; Prudhomme, D. R.; Buck, J. R.; Park, M.; Rizzo, C. J. Stereocontrolled syntheses of deoxyribonucleosides via photoinduced electron-transfer deoxygenation of benzoyl-protected ribo- and arabinonucleosides. *J. Org. Chem.* **2000**, *65*, 5969.

[149] Wang, Z. W.; Rizzo, C. J. Regioselective synthesis of beta-N1-and beta-N3-alloxazine nucleosides. *Org. Lett.* **2000**, *2*, 227.

[150] Kazimierczuk, Z.; Cottam, H. B.; Revankar, G. R.; Robins, R. K. Synthesis of 2'-Deoxytubercidin, 2'-Deoxyadenosine, and Related 2'-Deoxynucleosides Via a Novel Direct Stereospecific Sodium-Salt Glycosylation Procedure. *J. Am. Chem. Soc.* **1984**, *106*, 6379.

[151] Arico, J. W.; Calhoun, A. K.; McLaughlin, L. W. Preparation of the 2 '-Deoxynucleosides of 2,6-Diaminopurine and Isoguanine by Direct Glycosylation. *J. Org. Chem.* **2010**, *75*, 1360.

[152] Kotick, M. P.; Szantay, C.; Bardos, T. J. Synthesis of 5-S-Substituted 2'-Deoxyuridines . Study of Factors Influencing Stereoselectivity of Silyl Modification of Hilbert-Johnson Reaction. *J. Org. Chem.* **1969**, *34*, 3806.

[153] Rolland, V.; Kotera, M.; Lhomme, J. Convenient preparation of 2-deoxy-3,5-di-O-p-toluoyl-alpha-D-erythro-pentofuranosyl chloride. *Synth. Commun.* **1997**, *27*, 3505.

[154] Hubbard, A. J.; Jones, A. S.; Walker, R. T. An investigation by 1H NMR spectroscopy into the factors determining the beta:alpha ratio of the product in 2'-deoxynucleoside synthesis. *Nucleic Acids Res.* **1984**, *12*, 6827.

[155] Michel, J.; Gueguen, G.; Vercauteren, J.; Moreau, S. Triplex stability of oligodeoxynucleotides containing substituted quinazoline-2,4-(1H,3H)-dione. *Tetrahedron* **1997**, *53*, 8457.

[156] Godde, F.; Toulme, J. J.; Moreau, S. Benzoquinazoline derivatives as substitutes for thymine in nucleic acid complexes. Use of fluorescence emission of benzo[g]quinazoline-2,4-(1H,3H)-dione in probing duplex and triplex formation. *Biochemistry* **1998**, *37*, 13765.

[157] Hurley, L. H.; Kendrick, S.; Akiyama, Y.; Hecht, S. M. The i-Motif in the bcl-2 P1 Promoter Forms an Unexpectedly Stable Structure with a Unique 8:5:7 Loop Folding Pattern. *J. Am. Chem. Soc.* **2009**, *131*, 17667.

[158] Tor, Y.; Shin, D. W. Bifacial Nucleoside as a Surrogate for Both T and A in Duplex DNA. *J. Am. Chem. Soc.* **2011**, *133*, 6926.

[159] Pan, M. Y.; Hang, W.; Zhao, X. J.; Zhao, H.; Deng, P. C.; Xing, Z. H.; Qing, Y.; He, Y. Janus-type AT nucleosides: synthesis, solid and solution state structures. *Org. Biomol. Chem.* **2011**, *9*, 5692.

[160] Sandin, P.; Lincoln, P.; Brown, T.; Wilhelmsson, L. M. Synthesis and oligonucleotide incorporation of fluorescent cytosine analogue tC: a promising nucleic acid probe. *Nat. Protoc.* **2007**, *2*, 615.

[161] *Handbook of Chemical Glycosylation: Advances in Stereoselectivity and Therapeutic Relevance*; Demchenko, A. V., Ed.; Wiley-VCH Verlag GmbH & Co, 2008.

[162] Toshima, K.; Tatsuta, K. Recent Progress in O-Glycosylation Methods and Its Application to Natural-Products Synthesis. *Chem. Rev.* **1993**, *93*, 1503.

[163] Schmidt, R. R. New Methods for the Synthesis of Glycosides and Oligosaccharides - Are There Alternatives to the Koenigs-Knorr Method. *Angew. Chem., Int. Ed.* **1986**, *25*, 212.

[164] van der Marel, G. A.; Codee, J. D. C.; Litjens, R. E. J. N.; van den Bos, L. J.; Overkleeft, H. S. Thioglycosides in sequential glycosylation strategies. *Chem. Soc. Rev.* **2005**, *34*, 769.

[165] Crich, D.; Smith, M. 1-Benzenesulfinyl piperidine/trifluoromethanesulfonic anhydride: A potent combination of shelf-stable reagents for the low-temperature conversion of thioglycosides to glycosyl triflates and for the formation of diverse glycosidic linkages. *J. Am. Chem. Soc.* **2001**, *123*, 9015.

[166] Crich, D.; Pedersen, C. M.; Bowers, A. A.; Wink, D. J. On the use of 3,5-O-benzylidene and 3,5-O-(di-tert-butylsilylene)-2-O-benzylarabinothiofuranosides and their sulfoxides as glycosyl donors for the synthesis of beta-arabinofuranosides: Importance of the activation method. *J. Org. Chem.* **2007**, *72*, 1553.

[167] Young, R. J.; Shawponter, S.; Hardy, G. W.; Mills, G. Beta-Anomer Selectivity in 2'-Deoxynucleoside Synthesis - a Novel-Approach Using an Acyl Carbamate Directing Group. *Tetrahedron Lett.* **1994**, *35*, 8687.

[168] Hashimoto, S.; Inagaki, J.; Sakamoto, H.; Sano, A.; Nakajima, M. N-glycosylation with glycosyl diethyl phosphites: A highly stereoselective synthesis of 2 '-deoxy-beta-ribonucleosides. *Heterocycles* **1997**, *46*, 215.

[169] Sujino, K.; Sugimura, H. Stereocontrolled Synthesis of Beta-D-2'-Deoxyribonucleosides by Intramolecular Glycosylation. *Chem. Lett.* **1993**, 1187.

[170] Sandin, P.; Wilhelmsson, L. M.; Lincoln, P.; Powers, V. E. C.; Brown, T.; Albinsson, B. Fluorescent properties of DNA base analogue tC upon incorporation into DNA - negligible influence of neighbouring bases on fluorescence quantum yield. *Nucleic Acids Res.* **2005**, *33*, 5019.

[171] Vorbruggen, H.; Krolikiewicz, K.; Bennua, B. Nucleoside Syntheses .22. Nucleoside Synthesis with Trimethylsilyl Triflate and Perchlorate as Catalysts. *Chemische Berichte-Recueil* **1981**, *114*, 1234.

[172] Vorbruggen, H. Adventures in silicon - Organic chemistry. *Acc. Chem. Res.* **1995**, *28*, 509.

[173] Niedball.U; Vorbrugg.H. Synthesis of Nucleosides .9. General Synthesis of N-Glycosides .1. Synthesis of Pyrimidine Nucleosides. *J. Org. Chem.* **1974**, *39*, 3654.

[174] Kurosu, M.; Li, K. Highly Efficient O-Glycosylations with p-Tolyl Thioribosides and p-TolSOTf. *J. Org. Chem.* **2008**, *73*, 9767.

[175] Liu, Z. G.; Li, D. Y.; Yin, B. L.; Zhang, J. C. Highly stereoselective synthesis of 2 '-deoxy-beta-ribonucleosides via a 3 '-(N-acetyl)-glycyl-directing group. *Tet. Lett.* **2010**, *51*, 240.

[176] Ishiwata, A.; Akao, H.; Ito, Y. Stereoselective synthesis of a fragment of mycobacterial arabinan. *Org. Lett.* **2006**, *8*, 5525.

[177] During optimization of these conditions, variable temperatures and activators were tested. The use of a weaker activating agent such as NBS resulted in lower yields and increased reaction times. A temperature of 0 ˚C was found to be optimal, because lower temperatures resulted in incomplete conversion, and higher temperatures resulted in lower b-selectivity.

[178] Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraserreid, B. Armed and Disarmed N-Pentenyl Glycosides in Saccharide Couplings Leading to Oligosaccharides. *J. Am. Chem. Soc.* **1988**, *110*, 5583.

[179] Sandin, P.; Wilhelmsson, M.; Lincoln, P.; Norden, B.; Albinsson, B. tC - The first example of a non perturbing DNA base analogue with high, stable fluorescence quantum yield in nucleic acid systems. *Biophys. J.* **2005**, *88*, 410a.

[180] Motawi, M. S.; Marcussen, J.; Moller, B. L. A general method based on the use of N-bromosuccinimide for removal of the thiophenyl group at the anomeric position to generate a reducing sugar with the original protecting groups still present. *J. Carbohydr. Chem.* **1995**, *14*, 1279.

[181] Williams, D. R.; Klingler, F. D.; Allen, E. E.; Lichtenthaler, F. W. Bromine as an Oxidant for Direct Conversion of Aldehydes to Esters. *Tetrahedron Lett.* **1988**, *29*, 5087.

[182] Mandal, P. K.; Misra, A. K. Mild and efficient hydrolysis of thioglycosides to glycosyl hemiacetals using N-iodosaccharin. *Synlett* **2007**, 1207.

[183] Dinkelaar, J.; Witte, M. D.; van den Bos, L. J.; Overkleeft, H. S.; van der Marel, G. A. NIS/TFA: a general method for hydrolyzing thioglycosides. *Carbohydr. Res.* **2006**, *341*, 1723.

[184] Larsen, C. H.; Ridgway, B. H.; Shaw, J. T.; Woerpel, K. A. A stereoelectronic model to explain the highly stereoselective reactions of nucleophiles with five-membered-ring oxocarbenium ions. *J. Am. Chem. Soc.* **1999**, *121*, 12208.

[185] Larsen, C. H.; Ridgway, B. H.; Shaw, J. T.; Smith, D. M.; Woerpel, K. A. Stereoselective C-glycosylation reactions of ribose derivatives: Electronic effects of five-membered ring oxocarbenium ions. *J. Am. Chem. Soc.* **2005**, *127*, 10879.

[186] Smith, D. M.; Tran, M. B.; Woerpel, K. A. Nucleophilic additions to fused bicyclic five-membered ring oxocarbenium ions: Evidence for preferential attack on the inside face. *J. Am. Chem. Soc.* **2003**, *125*, 14149.

[187] Lavinda, O.; Tran, V. T.; Woerpel, K. A. Effect of conformational rigidity on the stereoselectivity of nucleophilic additions to five-membered ring bicyclic oxocarbenium ion intermediates. *Org. Biomol. Chem.* **2014**, *12*, 7083.

[188] Lavis, L. D.; Raines, R. T. Bright ideas for chemical biology. *ACS Chem. Biol.* **2008**, *3*, 142.

[189] Sletten, E. M.; Bertozzi, C. R. From Mechanism to Mouse: A Tale of Two Bioorthogonal Reactions. *Acc. Chem. Res.* **2011**, *44*, 666.

[190] Smith, L. M.; Sanders, J. Z.; Kaiser, R. J.; Hughes, P.; Dodd, C.; Connell, C. R.; Heiner, C.; Kent, S. B. H.; Hood, L. E. Fluorescence Detection in Automated DNA-Sequence Analysis. *Nature* **1986**, *321*, 674.

[191] Bao, G.; Rhee, W. J.; Tsourkas, A. Fluorescent Probes for Live-Cell RNA Detection. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 25.

[192] Srivatsan, S. G.; Greco, N. J.; Tor, Y. A highly emissive fluorescent nucleoside that signals the activity of toxic ribosome-inactivating proteins. *Angew. Chem., Int. Ed.* **2008**, *47*, 6661.

[193] Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Polarity of major grooves explored by using an isosteric emissive nucleoside. *ChemBioChem* **2008**, *9*, 706.

[194] Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. Emissive nucleosides as molecular rotors. *ChemPhysChem* **2011**, *12*, 567.

[195] Sun, K. M.; McLaughlin, C. K.; Lantero, D. R.; Manderville, R. A. Biomarkers for phenol carcinogen exposure act as pH-sensing fluorescent probes. *J. Am. Chem. Soc.* **2007**, *129*, 1894.

[196] Riedl, J.; Pohl, R.; Rulisek, L.; Hocek, M. Synthesis and photophysical properties of biaryl-substituted nucleos(t)ides. Polymerase synthesis of DNA probes bearing solvatochromic and pH-sensitive dual fluorescent and ^{19}F NMR labels. *J. Org. Chem.* **2012**, *77*, 1026.

[197] Teo, Y. N.; Kool, E. T. DNA-Multichromophore Systems. *Chem. Rev.* **2012**, *112*, 4221.

[198] Malinovskii, V. L.; Wenger, D.; Haner, R. Nucleic acid-guided assembly of aromatic chromophores. *Chem. Soc. Rev.* **2010**, *39*, 410.

[199] Brotschi, C.; Leumann, C. J. DNA with hydrophobic base substitutes: A stable, zipperlike recognition motif based on interstrand-stacking interactions. *Angew. Chem., Int. Ed.* **2003**, *42*, 1655.

[200] Zahn, A.; Leumann, C. J. Recognition properties of donor- and acceptor-modified biphenyl-DNA. *Chem. Eur. J.* **2008**, *14*, 1087.

[201] Takei, Y.; Yamaguchi, T.; Osamura, Y.; Fuke, K.; Kaya, K. Electronic-Spectra and Molecular-Structure of Biphenyl and Para-Substituted Biphenyls in a Supersonic Jet. *J. Phys. Chem.* **1988**, *92*, 577.

[202] Maus, M.; Rettig, W.; Bonafoux, D.; Lapouyade, R. Photoinduced intramolecular charge transfer in a series of differently twisted donor - Acceptor biphenyls as revealed by fluorescence. *J. Phys. Chem. A* **1999**, *103*, 3388.

[203] Maus, M.; Rettig, W.; Jonusauskas, G.; Lapouyade, R.; Rulliere, C. Subpicosecond transient absorption of donor-acceptor biphenyls. Intramolecular control of the excited state charge transfer processes by a weak electronic coupling. *J. Phys. Chem. A* **1998**, *102*, 7393.

[204] Maus, M.; Rettig, W. Electronic relaxations in donor-acceptor biphenyls. *Chem. Phys. Lett.* **2000**, *324*, 57.

[205] Rotkiewi, K.; Grellman, K. H.; Grabowsk, Z. R. Reinterpretation of Anomalous Fluorescence of "Para-N,N-Dimethylamino-Benzonitrile. *Chem. Phys. Lett.* **1973**, *19*, 315.

[206] Bhattacharyya, K.; Chowdhury, M. Environmental and Magnetic-Field Effects on Exciplex and Twisted Charge-Transfer Emission. *Chem. Rev.* **1993**, *93*, 507.

[207] Xie, Y.; Maxson, T.; Tor, Y. Fluorescent nucleoside analogue displays enhanced emission upon pairing with guanine. *Org. Biomol. Chem.* **2010**, *8*, 5053.

[208] Mata, G.; Luedtke, N. W. Stereoselective N-glycosylation of 2-deoxythioribosides for fluorescent nucleoside synthesis. *J. Org. Chem.* **2012**, *77*, 9006.

[209] Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y. Palladium-assisted routes to nucleosides. *Chem. Rev.* **2003**, *103*, 1875.

[210] Kinzel, T.; Zhang, Y.; Buchwald, S. L. A new palladium precatalyst allows for the fast Suzuki-Miyaura coupling reactions of unstable polyfluorophenyl and 2-heteroaryl boronic acids. *J. Am. Chem. Soc.* **2010**, *132*, 14073.

[211] Thakur, A.; Zhang, K.; Louie, J. Suzuki-Miyaura coupling of heteroaryl boronic acids and vinyl chlorides. *Chem. Commun.* **2012**, *48*, 203.

[212] Mata, G.; Luedtke, N. W. Synthesis and Solvatochromic Fluorescence of Biaryl Pyrimidine Nucleosides. *Org. Lett.* **2013**, *15*, 2462.

[213] Cavaluzzi, M. J.; Borer, P. N. Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res.* **2004**, *32*.

[214] Berlman, I. B. Handbook of Fluorescence Spectra of Aromatic Molecules. *Academic Press* **1971**.

[215] Rusakowi.R; Testa, A. C. 2-Aminopyridine as a Standard for Low-Wavelength Spectrofluorimetry. *J. Phys. Chem.* **1968**, *72*, 2680.

[216] Rettig, W. Charge Separation in Excited-States of Decoupled Systems - TICT Compounds and Implications Regarding the Development of New Laser-Dyes and the Primary Processes of Vision and Photosynthesis. *Angew. Chem., Int. Ed.* **1986**, *25*, 971.

[217] Grabowski, Z. R.; Dobkowski, J. Twisted Intramolecular Charge-Transfer (TICT) Excited-States - Energy and Molecular-Structure. *Pure Appl. Chem.* **1983**, *55*, 245.

[218] Ghoneim, N.; Suppan, P. Solvation of Tict States in Solvent Mixtures. *Pure Appl. Chem.* **1993**, *65*, 1739.

[219] Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994**, *94*, 2319.

[220] Sinkeldam, R. W.; Tor, Y. To D or not to D? On estimating the microenvironment polarity of biomolecular cavities. *Org. Biomol. Chem.* **2007**, *5*, 2523.

[221] Deng, F.; Kubin, J.; Testa, A. C. Deuterium isotope effects on the fluorescence of phenylpyridines. *Journal of Photochemistry and Photobiology a-Chemistry* **1997**, *104*, 65.

[222] Mayr, H.; Lakhdar, S.; Maji, B.; Ofial, A. R. A quantitative approach to nucleophilic organocatalysis. *Beilstein J. Org. Chem.* **2012**, *8*, 1458.

[223] Mayr, H.; Patz, M. Scales of Nucleophilicity and Electrophilicity - a System for Ordering Polar Organic and Organometallic Reactions. *Angewandte Chemie-International Edition in English* **1994**, *33*, 938.

[224] Ammer, J.; Nolte, C.; Mayr, H. Free Energy Relationships for Reactions of Substituted Benzhydryliunn Ions: From Enthalpy over Entropy to Diffusion Control. *J. Am. Chem. Soc.* **2012**, *134*, 13902.

[225] Mayr, H.; Ofial, A. R. Do general nucleophilicity scales exist? *J. Phys. Org. Chem.* **2008**, *21*, 584.

[226] Mayr, H.; Kempf, B.; Ofial, A. R. pi-nucleophilicity in carbon-carbon bond-forming reactions. *Acc. Chem. Res.* **2003**, *36*, 66.

[227] Mayr, H.; Bug, T.; Gotta, M. F.; Hering, N.; Irrgang, B.; Janker, B.; Kempf, B.; Loos, R.; Ofial, A. R.; Remennikov, G.; Schimmel, H. Reference scales for the characterization of cationic electrophiles and neutral nucleophiles. *J. Am. Chem. Soc.* **2001**, *123*, 9500.

[228] Gotta, M. F.; Mayr, H. Kinetics of the Friedel-Crafts alkylations of heterocyclic arenes: Comparison of the nucleophilic reactivities of aromatic and nonaromatic pi-systems. *J. Org. Chem.* **1998**, *63*, 9769.

[229] Netzel, T. L.; Zhao, M.; Nafisi, K.; Headrick, J.; Sigman, M. S.; Eaton, B. E. Photophysics of 2'-Deoxyuridine (Du) Nucleosides Covalently Substituted with Either 1-Pyrenyl or 1-Pyrenoyl - Observation of Pyrene-to-Nucleoside Charge-Transfer Emission in 5-(1-Pyrenyl)-Du. *J. Am. Chem. Soc.* **1995**, *117*, 9119.

[230] Baker-Austin, C.; Dopson, M. Life in acid: pH homelostasis in acidophiles. *Trends Microbiol.* **2007**, *15*, 165.

[231] Robinson, H.; Wang, A. H. J. 5'-Cga Sequence Is a Strong Motif for Homo Base-Paired Parallel-Stranded DNA Duplex as Revealed by Nmr Analysis. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5224.

[232] Volker, J.; Klump, H. H. Electrostatic Effects in DNA Triple Helices. *Biochemistry* **1994**, *33*, 13502.

[233] Plum, G. E.; Breslauer, K. J. Thermodynamics of an Intramolecular DNA Triple-Helix - a Calorimetric and Spectroscopic Study of the Ph and Salt Dependence of Thermally-Induced Structural Transitions. *J. Mol. Biol.* **1995**, *248*, 679.

[234] Gaffney, B. L.; Kung, P. P.; Wang, C.; Jones, R. A. Nitrogen-15-Labeled Oligodeoxynucleotides .8. Use of ^{15}N- NMR to Probe Hoogsteen Hydrogen-Bonding at Guanine and Adenine N7 Atoms of a DNA Triplex. *J. Am. Chem. Soc.* **1995**, *117*, 12281.

[235] Yu, A.; Barren, M. D.; Romero, R. M.; Christy, M.; Gold, B.; Dai, J. L.; Gray, D. M.; Haworth, I. S.; Mitas, M. At physiological pH, d(CCG)(15) forms a hairpin containing protonated cytosines and a distorted helix. *Biochemistry* **1997**, *36*, 3687.

[236] Asensio, J. L.; Lane, A. N.; Dhesi, J.; Bergqvist, S.; Brown, T. The contribution of cytosine protonation to the stability of parallel DNA triple helices. *J. Mol. Biol.* **1998**, *275*, 811.

[237] Nixon, P. L.; Giedroc, D. P. Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J. Mol. Biol.* **2000**, *296*, 659.

[238] Das, S. R.; Piccirilli, J. A. General acid catalysis by the hepatitis delta virus ribozyme. *Nat. Chem. Biol.* **2005**, *1*, 45.

[239] Nikolova, E. N.; Goh, G. B.; Brooks, C. L.; Al-Hashimi, H. M. Characterizing the Protonation State of Cytosine in Transient G·C Hoogsteen Base Pairs in Duplex DNA. *J. Am. Chem. Soc.* **2013**, *135*, 6766.

[240] Knitt, D. S.; Narlikar, G. J.; Herschlag, D. Dissection of the Role of the Conserved G-·-U Pair in Group-I Rna Self-Splicing. *Biochemistry* **1994**, *33*, 13864.

[241] Dai, J. X.; Hatzakis, E.; Hurley, L. H.; Yang, D. Z. I-Motif Structures Formed in the Human *c-MYC* Promoter Are Highly Dynamic-Insights into Sequence Redundancy and I-Motif Stability. *PLoS One* **2010**, *5*.

[242] Nesterova, I. V.; Nesterov, E. E. Rational Design of Highly Responsive pH Sensors Based on DNA i-Motif. *J. Am. Chem. Soc.* **2014**, *136*, 8843.

[243] Zhou, J.; Wei, C. Y.; Jia, G. Q.; Wang, X. L.; Feng, Z. C.; Li, C. Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. BioSyst.* **2010**, *6*, 580.

[244] Li, W.; Miyoshi, D.; Nakano, S.; Sugimoto, N. Structural competition involving G-quadruplex DNA and its complement. *Biochemistry* **2003**, *42*, 11736.

[245] Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Pluckthun, A. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8572.

[246] Phan, A. T.; Gueron, M.; Leroy, J. L. The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* **2000**, *299*, 123.

[247] Gueron, M.; Leroy, J. L. The i-motif in nucleic acids. *Curr. Opin. Struct. Biol.* **2000**, *10*, 326.

[248] Du, Z. H.; Yu, J. H.; Chen, Y. H.; Andino, R.; James, T. L. Specific recognition of the C-rich strand of human telomeric DNA and the RNA template of human telomerase by the first KH domain of human poly(C)-binding protein-2. *J. Biol. Chem.* **2004**, *279*, 48126.

[249] Mergny, J. L. Fluorescence energy transfer as a probe for tetraplex formation: The i-motif. *Biochemistry* **1999**, *38*, 1573.

[250] Nadler, A.; Strohmeier, J.; Diederichsen, U. 8-Vinyl-2 '-deoxyguanosine as a Fluorescent 2 '-Deoxyguanosine Mimic for Investigating DNA Hybridization and Topology. *Angew. Chem., Int. Ed.* **2011**, *50*, 5392.

[251] Tanpure, A. A.; Pawar, M. G.; Srivatsan, S. G. Fluorescent Nucleoside Analogs: Probes for Investigating Nucleic Acid Structure and Function. *Isr. J. Chem.* **2013**, *53*, 366.

[252] Liu, C. H.; Martin, C. T. Fluorescence characterization of the transcription bubble in elongation complexes of T7 RNA polymerase. *J. Mol. Biol.* **2001**, *308*, 465.

[253] Akiyama, Y.; Ma, Q.; Edgar, E.; Laikhter, A.; Hecht, S. M. A novel DNA hairpin substrate for bleomycin. *Org. Lett.* **2008**, *10*, 2127.

[254] Choi, J.; Tanaka, A.; Cho, D. W.; Fujitsuka, M.; Majima, T. Efficient electron transfer in i-motif DNA with a tetraplex structure. *Angew. Chem.* **2013**, *52*, 12937.

[255] Park, J. W.; Seo, Y. J.; Kim, B. H. Fluorescence modification of the AAAA (4A) loop: toward a probe of the structural dynamics of the i-motif of the retinoblastoma gene. *Chem. Commun.* **2014**, *50*, 52.

[256] Liu, D.; Zhang, Z. Y.; Zhang, H. Y.; Wang, Y. A novel approach towards white photoluminescence and electroluminescence by controlled protonation of a blue fluorophore. *Chem. Commun.* **2013**, *49*, 10001.

[257] Achelle, S.; Rodriguez-Lopez, J.; Robin-le Guen, F. Synthesis and Photophysical Studies of a Series of Quinazoline Chromophores. *J. Org. Chem.* **2014**.

[258] Surry, D. S.; Buchwald, S. L. Biaryl phosphane ligands in palladium-catalyzed amination. *Angew. Chem., Int. Ed.* **2008**, *47*, 6338.

[259] Melhuish, W. H. Quantum Efficiencies of Fluorescence of Organic Substances - Effect of Solvent and Concentration of Fluorescent Solute. *J. Phys. Chem.* **1961**, *65*, 229.

[260] Butler, R. S.; Cohn, P.; Tenzel, P.; Abboud, K. A.; Castellano, R. K. Synthesis, Photophysical Behavior, and Electronic Structure of Push-Pull Purines. *J. Am. Chem. Soc.* **2009**, *131*, 623.

[261] Gray, D. M.; Ratliff, R. L.; Vaughan, M. R. Circular-Dichroism Spectroscopy of DNA. *Methods Enzymol.* **1992**, *211*, 389.

[262] Wu, S.; Wang, X. Y.; Ye, X. D.; Zhang, G. Z. pH-Induced Conformational Change and Dimerization of DNA Chains Investigated by Analytical Ultracentrifugation. *J. Phys. Chem. B* **2013**, *117*, 11541.

[263] Although the fluorescence emission spectra are shown with λex = 260 nm, site-selective excitation of the probe at λex = 365 – 405 nm were also performed to calculate the quantum yields of DMAC in DNA.

[264] Vaya, I.; Gustavsson, T.; Miannay, F. A.; Douki, T.; Markovitsi, D. Fluorescence of Natural DNA: From the Femtosecond to the Nanosecond Time Scales. *J. Am. Chem. Soc.* **2010**, *132*, 11834.

[265] Kelley, S. O.; Barton, J. K. Electron transfer between bases in double helical DNA. *Science* **1999**, *283*, 375.

[266] Xu, D. G.; Nordlund, T. M. Sequence dependence of energy transfer in DNA oligonucleotides. *Biophys. J.* **2000**, *78*, 1042.

[267] Nordlund, T. M. Sequence, structure and energy transfer in DNA. *Photochem. Photobiol.* **2007**, *83*, 625.

[268] Zhang, D. Y.; Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **2009**, *131*, 17303.

[269] Genot, A. J.; Zhang, D. Y.; Bath, J.; Turberfield, A. J. Remote Toehold: A Mechanism for Flexible Control of DNA Hybridization Kinetics. *J. Am. Chem. Soc.* **2011**, *133*, 2177.

[270] Tang, W.; Wang, H.; Wang, D.; Zhao, Y.; Li, N.; Liu, F. DNA tetraplexes-based toehold activation for controllable DNA strand displacement reactions. *J. Am. Chem. Soc.* **2013**, *135*, 13628.

[271] Zhang, D. Y.; Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chem.* **2011**, *3*, 103.

[272] Chen, C.; Li, M.; Xing, Y. Z.; Li, Y. M.; Joedecke, C. C.; Jin, J.; Yang, Z. Q.; Liu, D. S. Study of pH-Induced Folding and Unfolding Kinetics of the DNA i-Motif by Stopped-Flow Circular Dichroism. *Langmuir* **2012**, *28*, 17743.

[273] Dettler, J. M.; Buscaglia, R.; Cui, J. J.; Cashman, D.; Blynn, M.; Lewis, E. A. Biophysical Characterization of an Ensemble of Intramolecular i-Motifs Formed by the Human c-MYC NHE III1 P1 Promoter Mutant Sequence. *Biophys. J.* **2010**, *99*, 561.

[274] Studies have shown that the stability of G-C and A-T Watson-Crick base pairs is pH-independent over the broad range of pH = 2 – 10.

[275] Lipps, H. J.; Rhodes, D. G-quadruplex structures: in vivo evidence and function. *Trends in Cell Biology* **2009**, *19*, 414.

[276] Chen, Y.; Qu, K.; Zhao, C.; Wu, L.; Ren, J.; Wang, J.; Qu, X. Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat Commun* **2012**, *3*, 1074.

[277] Oganesian, L.; Karlseder, J. Mammalian 5 ' C-Rich Telomeric Overhangs Are a Mark of Recombination-Dependent Telomere Maintenance. *Mol. Cell* **2011**, *42*, 224.

[278] Oganesian, L.; Karlseder, J. 5 ' C-rich telomeric overhangs are an outcome of rapid telomere truncation events. *DNA Repair* **2013**, *12*, 238.

[279] Dhakal, S.; Yu, Z. B.; Konik, R.; Cui, Y. X.; Koirala, D.; Mao, H. B. G-Quadruplex and i-Motif Are Mutually Exclusive in ILPR Double-Stranded DNA. *Biophys. J.* **2012**, *102*, 2575.

[280] Stellwagen, E.; Stellwagen, N. C. Single-Stranded DNA Oligomers may have a Molten Globule-Like Conformation in Solution. *Biophys. J.* **2011**, *100*, 58.

[281] Zhou, J.; Wei, C. Y.; Jia, G. Q.; Wang, X. L.; Feng, Z. C.; Li, C. Formation and stabilization of G-quadruplex in nanosized water pools. *Chem. Commun.* **2010**, *46*, 1700.

[282] de Lange, T. T-loops and the origin of telomeres. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 323.

[283] Skourti-Stathaki, K.; Proudfoot, N. J. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.* **2014**, *28*, 1384.

[284] Dhakal, S.; Schonhoft, J. D.; Koirala, D.; Yu, Z. B.; Basu, S.; Mao, H. B. Coexistence of an ILPR i-Motif and a Partially Folded Structure with Comparable Mechanical Stability Revealed at the Single-Molecule Level. *J. Am. Chem. Soc.* **2010**, *132*, 8991.

[285] Chakravarti, A. Population genetics - making sense out of sequence. *Nat. Genet.* **1999**, *21*, 56.

[286] Brookes, A. J. The essence of SNPs. *Gene* **1999**, *234*, 177.

[287] Cooper, D. N.; Smith, B. A.; Cooke, H. J.; Niemann, S.; Schmidtke, J. An Estimate of Unique DNA-Sequence Heterozygosity in the Human Genome. *Hum. Genet.* **1985**, *69*, 201.

[288] Frazer, K. A.; Murray, S. S.; Schork, N. J.; Topol, E. J. Human genetic variation and its contribution to complex traits. *Nat. Rev. Genet.* **2009**, *10*, 241.

[289] Pastinen, T.; Hudson, T. J. Cis-acting regulatory variation in the human genome. *Science* **2004**, *306*, 647.

[290] Sara, H.; Kallioniemi, O.; Nees, M. A Decade of Cancer Gene Profiling: From Molecular Portraits to Molecular Function. *Cancer Gene Profiling: Methods and Protocols* **2010**, *576*, 61.

[291] Arking, D. E.; Chakravarti, A. Understanding cardiovascular disease through the lens of genome-wide association studies. *Trends Genet.* **2009**, *25*, 387.

[292] Nebert, D. W.; Zhang, G.; Vesell, E. S. From human genetics and genomics to pharmacogenetics and pharmacogenomics: Past lessons, future directions. *Drug Metabolism Reviews* **2008**, *40*, 187.

[293] Carlson, C. S.; Newman, T. L.; Nickerson, D. A. SNPing in the human genome. *Curr. Opin. Chem. Biol.* **2001**, *5*, 78.

[294] Nakatani, K. Chemistry challenges in SNP typing. *ChemBioChem* **2004**, *5*, 1623.

[295] Okamoto, A.; Saito, Y.; Saito, I. Design of base-discriminating fluorescent nucleosides. *Journal of Photochemistry and Photobiology C, Photochemistry Reviews* **2005**, *6*, 108.

[296] Tainaka, K.; Tanaka, K.; Ikeda, S.; Nishiza, K.; Unzai, T.; Fujiwara, Y.; Saito, I.; Okamoto, A. PRODAN-conjugated DNA: Synthesis and photochemical properties. *J. Am. Chem. Soc.* **2007**, *129*, 4776.

[297] Mayer-Enthart, E.; Wagenknecht, H. A. Structure-sensitive and self-assembled helical pyrene array based on DNA architecture. *Angew. Chem., Int. Ed.* **2006**, *45*, 3372.

[298] Muller, J. Functional metal ions in nucleic acids. *Metallomics* **2010**, *2*, 318.

[299] Katz, S. The Reversible Reaction of Sodium Thymonucleate and Mercuric Chloride. *J. Am. Chem. Soc.* **1952**, *74*, 2238.

[300] Cantoni, O.; Christie, N. T.; Swann, A.; Drath, D. B.; Costa, M. Mechanism of $HgCl_2$ cytotoxicity in cultured mammalian cells. *Mol. Pharmacol.* **1984**, *26*, 360.

[301] Ariza, M. E.; Williams, M. V. Lead and mercury mutagenesis: Type of mutation dependent upon metal concentration. *J. Biochem. Mol. Toxicol.* **1999**, *13*, 107.

[302] Ono, A.; Togashi, H. Highly selective oligonucleotide-based sensor for mercury(II) in aqueous solutions. *Angew. Chem., Int. Ed.* **2004**, *43*, 4300.

[303] Xue, X. J.; Wang, F.; Liu, X. G. One-step, room temperature, colorimetric detection of mercury (Hg^{2+}) using DNA/nanoparticle conjugates. *J. Am. Chem. Soc.* **2008**, *130*, 3244.

[304] Wang, J.; Liu, B. Highly sensitive and selective detection of Hg^{2+} in aqueous solution with mercury-specific DNA and Sybr Green I. *Chem. Commun.* **2008**, 4759.

[305] Liu, C. W.; Hsieh, Y. T.; Huang, C. C.; Lin, Z. H.; Chang, H. T. Detection of mercury(II) based on Hg^{2+}-DNA complexes inducing the aggregation of gold nanoparticles. *Chem. Commun.* **2008**, 2242.

[306] Dumas, A.; Luedtke, N. W. Site-Specific Control of N7-Metal Coordination in DNA by a Fluorescent Purine Derivative. *Chem.--Eur. J.* **2012**, *18*, 245.

[307] Okamoto, A.; Ochi, Y.; Saito, I. Modulation of base selectivity for a base-discriminating fluorescent nucleobase by addition of mercury ion. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4279.

[308] Uchiyama, T.; Miura, T.; Takeuchi, H.; Dairaku, T.; Komuro, T.; Kawamura, T.; Kondo, Y.; Benda, L.; Sychrovsky, V.; Bour, P.; Okamoto, I.; Ono, A.; Tanaka, Y. Raman spectroscopic detection of the T-Hg^{II}-T base pair and the ionic characteristics of mercury. *Nucleic Acids Res.* **2012**, *40*, 5766.

[309] Wilhelmsson, L. M.; Sandin, P.; Lincoln, P.; Brown, T. Synthesis and oligonucleotide incorporation of fluorescent cytosine analogue tC: a promising nucleic acid probe. *Nature Protocols* **2007**, *2*, 615.

[310] Ocejo, M.; Carrillo, L.; Vicario, J. L.; Badia, D.; Reyes, E. Role of Pseudoephedrine as Chiral Auxiliary in the "Acetate-Type" Aldol Reaction with Chiral Aldehydes; Asymmetric Synthesis of Highly Functionalized Chiral Building Blocks. *J. Org. Chem.* **2011**, *76*, 460.

[311] Cantor, C. R.; Warshaw, M. M.; Shapiro, H. Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* **1970**, *9*, 1059.

[312] Tataurov, A. V.; You, Y.; Owczarzy, R. Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys. Chem.* **2008**, *133*, 66.